Titre: Title:	Mean Field Game-Based Control of Dispersed Energy Storage Devices with Constrained Inputs
Auteur: Author:	Feng Li
Date:	2016
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Li, F. (2016). Mean Field Game-Based Control of Dispersed Energy Storage Devices with Constrained Inputs [Master's thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2435/

UNIVERSITÉ DE MONTRÉAL

MEAN FIELD GAME-BASED CONTROL OF DISPERSED ENERGY STORAGE DEVICES WITH CONSTRAINED INPUTS

FENG LI
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE ÉLECTRIQUE)
DÉCEMBRE 2016

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

MEAN FIELD GAME-BASED CONTROL OF DISPERSED ENERGY STORAGE DEVICES WITH CONSTRAINED INPUTS

présenté par: LI Feng
en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de:

M. SAYDY Lahcen, Ph. D, président
M. MALHAMÉ Roland P., Ph. D., membre et directeur de recherche
M. LE NY Jérôme, Ph. D., membre et codirecteur de recherche
M. ZHU Guchuan, Ph. D., membre

ACKNOWLEDGEMENTS

I am sincerely grateful to my supervisors Professor Roland Malhamé and Professor Jérôme Le Ny at Polytechnique Montréal for their continuous guidance and supervision throughout the research work. In particular, I would like to thank Professor Malhamé for introducing me to the challenging and interesting topic of mean field game theory. He is a knowledgeable advisor who provided a lot of technical guidance and explanations to many theoretical details for the thesis. I wish to thank Professor Le Ny for his many inspiring ideas and supports to work out several technical difficulties in the thesis. I find myself very fortunate to have the opportunity to work with great scholars such as Professor Malhamé and Professor Le Ny.

My special thanks to Dr. Arman Kizilkale who provided me help and many advices in starting the research work.

I am grateful for the financial supports provided by Bourse Maurice Brisson - BBA and funding from Natural Resources Canada.

Last but not least, I would like to thank my family for their support and patience. This thesis is my gift to them.

RÉSUMÉ

Avec plus de production d'énergie renouvelable (éoliennes, panneaux solaires) connectés au réseau électrique, il devient important d'équilibrer la charge et la production alors que la variabilité est un problème pour les ressources renouvelables. On a étudié les dispositifs de stockage d'énergie, en particulier ceux qui sont naturellement présents au réseau électrique (tels que les chauffe-eau électriques, les chauffages électriques, etc. dans les ménages) pourraient devenir des sources potentielles pour atténuer ce problème de variabilité.

Afin de minimiser le besoin de communication et les efforts de calcul, un mécanisme de contrôle décentralisé est développé dans cette recherche pour contrôler la trajectoire de champ moyen d'une grande population de chauffages électriques pour suivre (ou atteindre) une cible de température. Le contrôle optimal pour chaque chauffage ou joueur est localement calculé et respecte certaines contraintes. Le problème de contrôle est formulé comme un problème de commande linéaire quadratique (LQ) avec des entrées contraintes dans la théorie du jeu de champ moyen, en anglais, Mean Field Game (MFG). Le mécanisme de contrôle décentralisé est développé à partir de la solution d'équilibre du modèle, cette dernière est obtenue lorsque chaque joueur individuel choisit la meilleure réponse au champ moyen basé sur l'état global de la population. En choisissant les stratégies optimales, tous les joueurs reproduisissent collectivement le champ moyen de la population. En ce sens, la solution d'équilibre est un point fixe du système. Dans une configuration de MFG, les joueurs sont faiblement couplés, ce qui signifie qu'un joueur a une influence négligeable sur le système global tandis que le comportement de la population entière a un effet significatif sur chaque joueur individuel.

Dans la première partie, le problème est décrit par un système d'équations à champ moyen avec les entrées contraintes, en supposant que la distribution de la température initiale de la population est connue. Le système d'équations est composé d'un nombre de K d'équations aux dérivées partielles (EDPs) couplées;

la solution décrit la propagation de la distribution de la population sur un horizon de contrôle, qui est ensuite utilisée pour calculer l'information de champ moyenne. Sur la base de l'information du champ moyen calculée, chaque joueur individuel peut déterminer son contrôle optimal.

La deuxième partie décrit la politique de contrôle qui est adoptée par le joueur individuel sous les entrées contraintes. Comme il est difficile de calculer la solution de contrôle optimale exacte au problème de LQ restreint, une politique de commande est développé qui permet au plus une commutation de contrôles saturés à des contrôles insaturés, où des contrôles saturés, le cas échéant, se produisent uniquement au début de l'horizon de contrôle. Après une période de temps, en fonction de la condition initiale du joueur, les contrôles deviennent insaturés et par la suite on peut calculer la politique de contrôle optimale en boucle fermée comme dans un problème de LQ sans contraint sur les entrées. La condition suffisante pour une telle politique de commande de commutation unique est proposée et prouvée, et un algorithme numérique pour calculer le point fixe du système est donné.

ABSTRACT

With the increasing levels of renewable power generation (wind turbines, solar panels) connected to power grids, it is becoming important to balance the load and the generation while variability issue exists for renewable resources. It has been studied that energy storage devices, in particular those naturally present in the power system (such as electric water heaters, electric space heaters, etc. in households), can become potential sources to help mitigate such variability.

In order to minimize the communication bandwidth and computation efforts, a decentralized control mechanism is developed in this research to shape the aggregate load profile of a large population of electric space heaters in a power system. Under the control mechanism, the mean-field trajectory (temperature) of the load population is controlled to follow a target temperature, and the control input for each device is generated locally and must respect certain constraint. We formulate the problem as a linear quadratic (LQ) tracking problem with constrained inputs under the mean field game (MFG) settings. The decentralized mechanism is then derived based on the equilibrium solution to the formulated model. The equilibrium solution is obtained when each individual agent of the game chooses a best response to the so-called *mass effect* of the population via couplings of their individual dynamics and cost functions; when all agents choose the optimal strategies, collectively the mass effect should be replicated. In this sense, the equilibrium solution is a fixed point of the system, and can be mathematically characterized by a mean-field (MF) fixed point equation system. In such MFG setup, agents are weakly coupled, meaning that an agent has a negligible influence on the overall system while the mean-field behavior of the population has a significant effect on any individual agent.

In the first part, we formulate the problem by a constrained model, and propose the MF fixed point equation system under constrained controls. We assume that the initial temperature distribution of the population is known. The MF equation system consists of K sets of coupled partial differential equations

(PDE's), the solution to which describes the propagation of the probability distribution of the population, which is then used to compute the mean-field information. Individual agent can then determine the optimal control under constraints as a best response to the computed mean-field information.

As it is rather difficult to compute the exact optimal control solution to the constrained model, in the second part we implement a single switching best response control policy where saturated controls, if any, happen only at the beginning of the control horizon. After certain period of time, depending on individual's initial condition, controls become unsaturated and we can compute the optimal control policy thereafter as in an unconstrained LQ problem. We prove a sufficient condition under which such single switching control policy is indeed the best response control, and we propose a numerical algorithm to compute the fixed point of the constrained system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

MFG	Mean Field Game
LQG	Linear Quadratic Gaussian
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
HJB	Hamilton-Jacobi-Bellman equation
FP	Fokker-Planck equation
PMP	Pontryagin's Maximum Principle

LIST OF APPENDIXES

CHAPTER 1 INTRODUCTION

1.1 Description of the Problem Studied

With the increasing levels of renewable power generation (wind turbines, solar panels) connected to power grids, it is becoming important to plan for and optimize renewable power generation, so as to manage the peak load, increase the efficiency and improve the economics of power sources. However, the intermittent characteristics of renewable generation (which amounts to high variability), although partially predictable, lead in general to a non-optimized power generation system. When less power is generated from renewable sources due to lack of source availability, the power plant must make up the gap by utilizing conventional power sources such as coal or natural gas, or using bulk energy storage devices such as battery packs, compressed air, etc [Tarroja *et al.*, 2012]. However, these approaches may incur high operating costs, reduce the efficiency of power generation, and reverse the environmental advantages and the positive effects of the renewable generation. In this context, energy storage, in particular that associated with a large number of dispersed devices naturally present in the power system (such as electric water heaters, electric space heaters, refrigerators, etc.), constitutes a readily available alternative to help mitigate the variability of renewable power generation [Callaway, 2009] (see Figure 1.1). To accommodate the status of renewable generation, we can reduce the aggregate power consumption of a population of devices when the generation is low, and increase the consumption when the generation is at surplus. In order to utilize and control these dispersed devices to manage the load profile of the entire population, one may model such a large-scale dynamic system as a large-scale control / optimization problem, and develop a centralized control architecture whereby each device is individually managed as a best response to each of all other devices' control actions; however, this requires significant computation efforts and a large bandwidth to communicate with and monitor each device. It also raises privacy issue for data exchange between the central control authority

and users and/or among users.

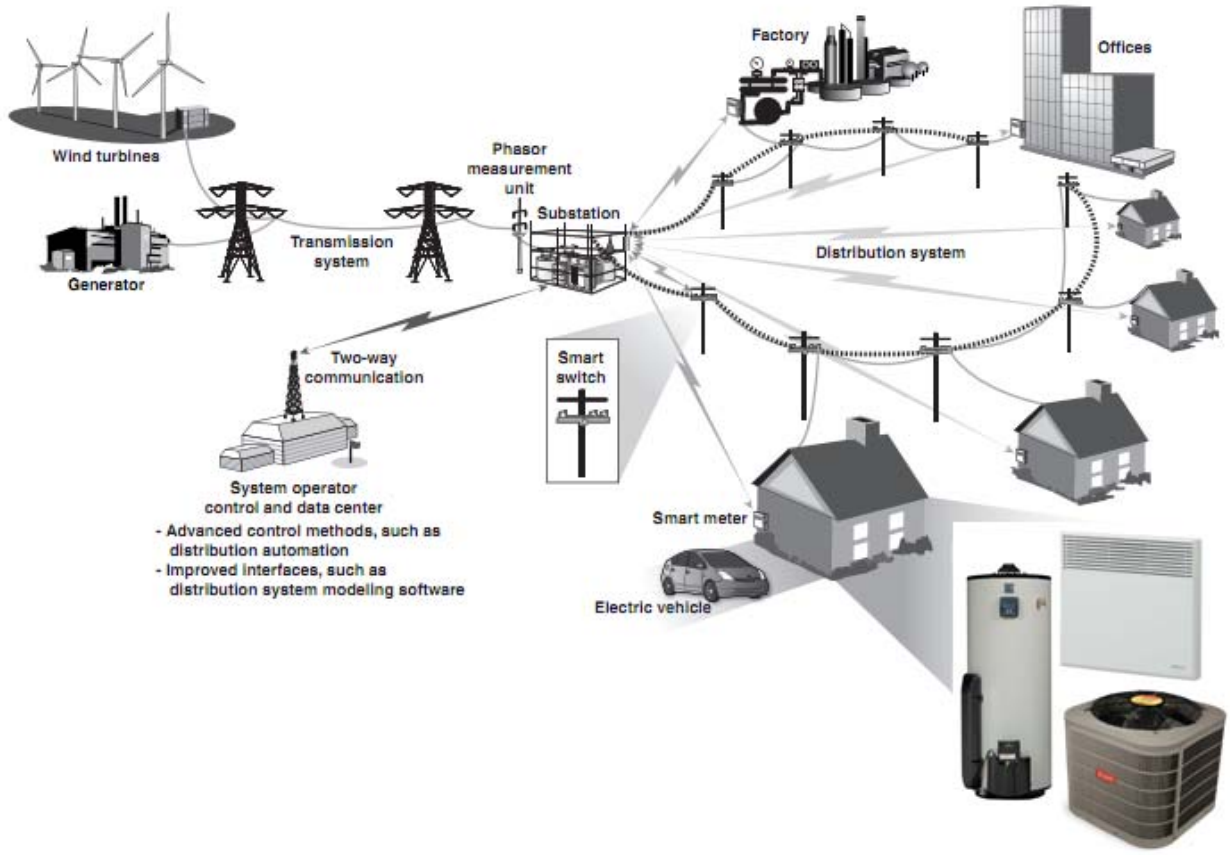

Figure 1.1 "Energy Internet" by including house appliances as energy storage devices. (*Source: Leonardo Energy*)

Given the challenges of centralized control methods in producing the desired results, a decentralized control mechanism needs to be developed and must meet the following requirements: (1) Control actions must be locally decided by each device yet preserve global optimality; (2) The level of data exchange between the central authority and users must be kept to a minimum; and (3) The disturbance to users' load profile relative to the uncontrolled situation must be kept to a minimum, and there should be no threat to privacy. Past researches formulate the coordination of a large number of energy storage devices (electric space heaters) in a power system under the linear quadratic Gaussian (LQG)

mean field game (MFG) setup, and lead to a class of decentralized control mechanisms that can actually satisfy these three requirements.

1.2 Research Objective

In this research, we extend the study of decentralized control of power systems with a large number of energy storage devices in [Kizilkale et Malhamé, 2013] under the MFG setup, but we consider *the case where the control inputs for the devices are constrained*. We are motivated by this subject as in most engineering applications control inputs may be constrained due to physical or design limitations of the system. When control input constraints are imposed, the system dynamics become non-linear where controls may become saturated, hence formulating the system under the LQG MFG setup fails in that, as it turns out, the optimal control solutions can no longer reduce to the solution of ordinary differential equations (ODE's). To overcome this difficulty, we propose to formulate this constrained MF control problem as a coupled system of partial differential equations (PDE's) [Huang *et al.*, 2006] to characterize the mean field effect, and solve the PDE's for a fixed point as the optimal control solutions.

The contributions of the research consist of the following:

- We present a MF fixed point equation system characterizing the limiting infinite population Nash equilibrium. The equation system describes the propagation of the probability distribution of the device population's temperatures and the optimal control solution under constraint for each device. We propose a numerical algorithm to find the fixed point solution to the equation system.
- We develop a sufficient condition under which any device's optimal control as a best response to the mean field involves either a single switching from saturated to unsaturated control, or a control which never saturates. We establish under this sufficient condition a numerical scheme to compute the best response controls of the devices.

- We argue that if all quantities converge to a steady-state equilibrium, the latter must be the identical in both the constrained and unconstrained cases. Thus it is only the transient dynamics which are affected by the saturation effects.

The thesis is thus organized as follows. Chapter 2 presents a literature review of models and approaches to solve dynamic systems with a large population under the mean field games (MFG) framework. Chapter 3 presents the MFG model and equation system to control the mean temperature of a large population of electric heaters under constrained control inputs. Each heater device must respect an input constraint when following its optimal control, and the mean temperature of the population must follow a set temperature target. In Chapter 4 we analyze the control policy, present our sufficient condition for at most single switching best response policies of individuals given a posited mean field, develop the fixed point equations characterizing the associated equilibrium mean field, and propose a numerical algorithm to search for the fixed point solution. In Chapter 5 we illustrate the equilibrium solution under constrained control inputs by numerical studies. Chapter 6 discusses the conclusion and proposes possible extensions to the subject of study.

CHAPTER 2 LITERATURE REVIEW

The concept of using household appliances as energy storage devices to adjust and shape the power demands of the loads on a power network is becoming an important demand management tool. One strategy is *direct load control* and it relies on developing aggregate models to describe the evolution of the mass probability distribution of a large population of thermostatically controlled loads (TCL) and to control their aggregate power demand to balance the load and generation. Different types of TCLs have been explored in the literature, for example, air conditioners [Callaway, 2009; Koch *et al.*, 2009] and water heaters [Gustafson *et al.*, 1993; Ericson, 2009]. In general, the aggregate load dynamics are often described by coupled PDE's [Laurent et Malhamé, 1994; Malhamé et Chong, 1985]. As the control of the aggregate load is mostly centralized in nature, efforts have been made to simplify the control architecture and facilitate extension to large scale systems. However, issues still remain in terms of communication bandwidth and requirement of state observation/estimation for heterogeneous groups of TCLs.

A class of decentralized control mechanisms for load management of power systems has been proposed in [Kizilkale et Malhamé, 2013, 2014b,a] based on the mean field game (MFG) equilibrium solution of linear quadratic Gaussian (LQG) models with integral control in the cost coefficients. For example, electric space heaters with diffusion dynamics are controlled to collectively reach a mean target temperature under a non-cooperative framework in [Kizilkale et Malhamé, 2013] and under a cooperative framework in [Kizilkale et Malhamé, 2014a]. In [Kizilkale et Malhamé, 2014b] water heaters with Markovian jump-driven hot water demand models are used as agents under the mean field control. The equilibrium solution concept used to derive these decentralized control mechanisms is based on the main results derived in [Huang *et al.*, 2007]. However, the cited literature does not consider constraints on control inputs when deriving optimal control laws. In contrast, [Bagagiolo et Bauso, 2013] and [Grammatico *et al.*, 2015] study controls of the power demands of a large

number of home appliances under an MFG framework, in the presence of saturated controls. A chattering switching control is implemented in [Bagagiolo et Bauso, 2013], and at equilibrium the mean temperature and the main frequency are regulated at the desired values. In contrast, we shall be dealing here with piecewise smooth controls. In [Grammatico *et al.*, 2015], a pricing incentives scheme is implemented for the entire population based on the current aggregated power demand where individual control is bang-bang-like switching. In [Paccagnan *et al.*, 2015], the set of feasible aggregated power trajectories are characterized such that a population of TCLs can follow them, where each TCL has a hybrid dynamics and the mode switching rate is constrained. The control is common to the entire population of TCLs and is announced by the central authority. The control is implemented as a saturation function such that it is proportional to the derivative of the desired power trajectory; when the derivative is too high, the control is saturated to some limiting value. In this implementation, the individual state of each TCL device should be known to the central authority at all times in order to compute the required controls.

In general, it is significantly more difficult to find the optimal solution to an input constrained infinite-horizon LQR problem than the unconstrained case, and often the solution is computed and approximated by numerical approaches. An approximation scheme which is widely studied is model predictive control (MPC), which reduces the infinite-time constrained problem into a finite-time one by choosing a prediction horizon, and an open-loop optimal solution can be approximated, for example in [Scokaert et Rawlings, 1998]. One disadvantage of this approach is that computations are performed online, so it may not be feasible for applications where short prediction horizon is required and many state variables are involved. Algorithms are therefore developed by formulating the constrained problem as a multiparametric quadratic program, and trying to find a feedback optimal solution that can be computed offline. For example, works by [Bemporad *et al.*, 2000; Seron *et al.*, 2000; Grieder *et al.*, 2003] study how a lookup table can be implemented by computing piecewise affine (PWA) solutions offline, while the online computation reduces to evaluation of the function. This significantly reduces the complexity of online computation

and implementation. The MPC approach, either online or offline, can be applied when there are constraints on the state variables as well.

Non-MPC methodologies are also explored to appoximate the optimal solution to the constrained LQR problem. In [Goebel et Subbotin, 2005], a control problem dual to the constrained LQR problem is studied, which has no constraint but non-quadratic cost. The value function is then described as a convex conjugate of the dual value function, and an optimal feedback solution is constructed based on the gradient of the optimal value function, which can be numerically propagated backwards from the solution to an appropriate Riccati equation. An algorithm is proposed to develop a lookup table of the optimal solution corresponding to all initial conditions at all times. The optimal solution for a given state is then matched to that obtained from the same initial condition. In [Hassan et Boukas, 2008], some necessary conditions for optimality are derived under which an optimal solution is a linear state-feedback with saturation boundaries. Then the optimal solution is computed from an iterative numerical approach based on correcting the error between the computed control at each iteration and the desired control trajectories satisfying the necessary conditions. In [Heemels *et al.*, 1998], the LQR problem with positive control constraint is studied. Although no numerical algorithms are proposed to find an optimal solution, some necessary and sufficient conditions for optimality are developed under which an optimal solution can be written as a feedback of the projection onto a closed convex cone of positive controls. The optimal solution under these conditions can then be solved from the Maximum Principle or based on the dynamic programming equation. In [Grammatico *et al.*, 2016], mean field problems under heterogeneous convex state and control constraints are considered , and several decentralized iteration approaches and conditions to converge to a fixed point asymptotic Nash equilibrium are proposed. The approach does not rely on optimal control theory, requires continuous communication with a central coordinator, and formulates the search for best response policies as a large size static optimization problem. Works by [Graichen et Petit, 2008; Graichen *et al.*, 2010] study methods to transform the inequalities on the state and input constraints into equalities through some saturation functions,

then construct an extended unconstrained system. The optimal solution to the extended system can then be numerically solved as a two-point boundary value problem (BVP), where optimality conditions can be derived from the associated Hamiltonian and the Maximum Principle.

CHAPTER 3 MEAN FIELD GAME MODEL WITH CONSTRAINED CONTROL INPUTS

The following notation is used throughout the entire document.

$x^i(t)$	Individul agent i's state (temperature)
x_a	Outside ambient temperature, assumed to be constant
ρ_0	Probability density function (pdf) of initial temperatures
x_0^i	Initial temperature of agent i, drawn randomly from ρ_0
$\bar{x}(t)$	Mean temperature of the population
u_{free}^i	Free control for agent i to stay at x_0^i
$u^i(t)$	Control which deviates agent i from its x_0^i
$u_{total}^i(t)$	Total control for agent i where $u_t^i(t) = u^i(t) + u_{free}^i$
$\mathcal{U}^i$	Control constraint which depends on i
y	Target mean temperature set by the central authority
z	Temperature boundary to maintain users' comfort levels

3.1 Mean Field Game Model with Constrained Control

We consider a large population of N space heating devices (agents of the game) in the power grid. It is assumed that a central authority (either the power system authority or an "aggregator" managing the group of devices) would like the mean state of all agents to follow a target trajectory. We assume that the dynamics of each agent $i \in [N] := \{1, \ldots, N\}$ are described by the following

process

$$\frac{dx^i}{dt} = -a(x^i - x_a) + bu^i_{total}, \;\; t \geq 0, \; x^i(0) = x^i_0, \tag{3.1}$$

where $a > 0$, $b > 0$, and x_a is the outside ambient temperature. Let $x^i_0 \in \mathbb{R}$ be agent i's random initial condition, distributed according to some known probability density function (pdf) ρ_0. We assume $x_a < x^i_0, \;\forall i \in [N]$. Note that in this model, the thermostat control of the space heater is replaced by a continuous control $u^i_{total}(t)$.

It is natural that each device will want to remain at its initial temperature x^i_0, which requires a control u^i_{free}. Such u^i_{free} is considered *free* in the sense that in the quadratic cost function defined below we do not penalize devices' control efforts to remain at their initial temperatures. Hence we are only interested in the control effort u^i which is required to deviate an agent i from x^i_0 and the contribution to moving the mean temperature of the population towards the target temperature y. Also for the rest of the thesis, we consider only the deterministic case where there is no noise in devices' dynamics. We can reformulate the dynamics in (3.1) by writing $u^i_{total} = u^i + u^i_{free}$

$$\frac{dx^i}{dt} = \{-a(x^i - x^i_0) + bu^i\} + \{-a(x^i_0 - x_a) + bu^i_{free}\}, \;\; t \geq 0, \; x^i(0) = x^i_0. \tag{3.2}$$

By the definition above of free control u^i_{free}, the latter term in (3.2) becomes zero, hence we have

$$u^i_{free} \triangleq b^{-1}a(x^i_0 - x_a)$$

which is constant and depends on the individual initial temperature. Recall that we assume that x_a is constant. We then have a simplified dynamics model

$$\frac{dx^i}{dt} = -a(x^i - x^i_0) + bu^i \triangleq f(x^i, u^i), \;\; t \geq 0, \; x^i(0) = x^i_0. \tag{3.3}$$

Under the control constraint, it is the total control $u^i_{tot} = u^i + u^i_{free}$ that is constrained to a common set of values $\mathcal{U}_{total} \triangleq \{u_{total} | u_{\min} \leq u_{total} \leq u_{\max}\}$ for all N agents. The constraints $u_{\min}$ and $u_{\max}$ need to be chosen carefully

in order to make sense physically. Specifically, we must have $0 \leq u_{\min} \leq \min\{u^i_{free}, \forall i \in [N]\}$, and $u_{\max} \geq \max\{u^i_{free}, \forall i \in [N]\}$. These are imposed because at $t = 0$ all u^i_{free} should fall within the admissible control set $[u_{\min}, u_{\max}]$. Also, the minimum control u^i_{total} that agent i can exert is zero heating, hence $u_{\min} \geq 0$. Hence we have the admissible control $u^i \in \mathcal{U}^i \triangleq \{u^i | u_{\min} - u^i_{free} \leq u^i \leq u_{\max} - u^i_{free}\}$, where $u_{\min}$ and $u_{\max}$ meet the conditions mentioned above. Note in particular that the constraint set for u^i depends on the agent i.

The goal is to move the average temperature of the agents' population to the target y, while keeping each individual temperature relatively close to its initial value. In particular, we want to avoid controls u^i to become too aggressive, in which case each individual temperature x^i could be moved to y. Although this satisfies our goal to move the population mean to the mean target, this will inevitably cause disturbance to users whose initial temperature is far from the mean target. This motivates the definition of the following cost function for each agent i, and we define u^{-i} to be the set of controls of agents other than agent i, i.e. $u^{-i} = (u^1, u^2, \cdots, u^{i-1}, u^{i+1}, \cdots, u^N)$.

$$J_i(u^i, u^{-i}) = \int_0^\infty e^{-\delta t}[(x^i - z)^2 q_t + (x^i - x^i_0)^2 q_0 + (u^i)^2 r]dt, \tag{3.4}$$

where δ , q_0 , r are positive constants and the temperature z serves as a direction signal to all agents, such that all agents should move toward z but not beyond. Deviation of temperature from x^i_0 is penalized by the coefficient q_0 in order to maintain users' comfort levels by keeping temperatures close to the initial values. The penalty coefficient q_t, defined below, is calculated according to the integrated difference between the mean field temperature of the entire population $\bar{x}$ and the constant target y

$$q_t = \left| \lambda \int_0^t (\bar{x} - y)dt \right| + k_p(t), \tag{3.5}$$

where $\lambda > 0$ and is constant. $\bar{x}$ is computed from the propagation of the mass density of the population, and is defined in (3.9).

The cost function J_i is defined so that any agent i feels the pressure q_t built up from the difference between the mean field temperature $\bar{x}$ of the

population and the target temperature y. In the case where $y < \bar{x}_0$, all agents are asked to reduce their temperatures, and one must have $\bar{x} \geq y > z > x_a$. If q_t is to achieve a steady-state, then $\bar{x}$ must approach y asymptotically. At this point, one can show that all agents reach an individualized, initial condition dependent steady-state and maintain the population mean temperature at y. $k_p(t)$ is a positive time-varying term, which provides some initial pressure to all agents at the beginning of the control horizon, and guarantees a maximum single switch property of the locally optimal control policy, which will be discussed in the following chapter. As q_t is calculated based on mean temperature $\bar{x}$, and the latter becomes deterministic as the number of agents goes to infinity, q_t can be viewed in the limit as a given function of time t, and we then write $J_i(u^i, q_t)$. For a posited q_t trajectory, each agent chooses its best response $u^i_* = \arg\min_{u^i \in \mathcal{U}^i} J_i(u^i, q_t)$ to minimize J_i while respecting the control constraints. By defining the value function $V_i(x^i, t)$ as the optimal cost to go starting from x^i at time t, we write the Hamilton-Jacobi-Bellman (HJB) equation for agent i to find the optimal u^i_* under constraint

$$-\frac{\partial V_i}{\partial t} = \inf_{u^i \in \mathcal{U}^i} \left(L(x^i, u^i, t) + \frac{\partial V_i}{\partial x^i} f(x^i, u^i) \right), \tag{3.6}$$

where

$$\begin{aligned} L(x^i, u^i, t) &= \frac{1}{2} e^{-\delta t} [(x^i - z)^2 q_t + (x^i - x^i_0)^2 q_0 + (u^i)^2 r], \\ f(x^i, u^i) &= -a(x^i - x^i_0) + b u^i. \end{aligned}$$

As in the usual MFG set up, we consider a limiting infinite population situation where the mass effect $\bar{x}$ characterizing q_t, the weight in the cost function of individual agents, is posited as given. The solution to the HJB equation describes the optimal control u^{i*} as a best response to the *mass* behavior for each agent i. However, *in contrast to the standard linear quadratic MFG setup* (More background information on standard LQG MFG and its equilibrium solution can be found in Annexe A.), the dynamics in (3.3), the cost function in (3.4) and the constraint set $\mathcal{U}^i$ depend throughout the control horizon on the initial condition x^i_0 of the particular agent. To address this difficulty and characterize

the mass behavior, we still view the population of agents as a continuum, but approximate the mass dynamics by discretizing the agents' initial temperature pdf ρ_0. First, we partition the set of initial temperatures into disjoint intervals $\Theta_k, k = 1, \ldots, K$. We define $m^{\Theta_k}(x,t)$ to be the temperature pdf *conditional* on the initial temperature falling in the set Θ_k. For Θ_k, $k = 1, \ldots, K$, the initial conditional pdf $m^{\Theta_k}(x,0)$ is supported and assumed uniform over the intervals Θ_k. Moreover, we attach to an interval Θ_k a single control policy, taken to be the one associated with the particle initially at the mean temperature $\bar{\theta}_k$ within Θ_k. In practice we take $\bar{\theta}_k$ as the center of the interval Θ_k, where m^{Θ_k} is uniform.

We can then compute $m^{\Theta_k}(x,t)$ using the advection equation

$$\frac{\partial m^{\Theta_k}(x,t)}{\partial t} + \frac{\partial}{\partial x}\{m^{\Theta_k} v^{\bar{\theta}_k}\} = 0, \tag{3.7}$$

where $v^{\bar{\theta}_k} := f_k(x, u_*^{\bar{\theta}_k}) = -a(x - \bar{\theta}_k) + b u_*^{\bar{\theta}_k}$.

The optimal control $u_*^{\bar{\theta}_k}$ is obtained from the HJB equation

$$-\frac{\partial V_k}{\partial t} = \inf_{u \in \mathcal{U}^{\bar{\theta}_k}} \left(L_k(x,u,t) + \frac{\partial V_k}{\partial x} f_k(x,u) \right). \tag{3.8}$$

Finally, in order to calculate the mean $\bar{x}$ of the mass, we have

$$\bar{x} = \int_{-\infty}^{+\infty} x m(x,t) dx, \tag{3.9}$$

where the mean-field $m(x,t)$ is computed as

$$m(x,t) = \sum_{k=1}^{K} m^{\Theta_k}(x,t) \left(\int_{\Theta_k} \rho_0(\alpha) d\alpha \right).$$

The MFG equation system to solve consists of (3.5), (3.7), (3.8) and (3.9), with $k = 1, \ldots, K$. As discussed below, it corresponds to requiring that at equilibrium in an infinite population, a fixed point property be satisfied.

MF Equation System with Constrained Control for $t \in [0, +\infty)$

$$\begin{aligned}
\frac{\partial m^{\Theta_k}(x,t)}{\partial t} + \frac{\partial}{\partial x}\{m^{\Theta_k} v^{\bar{\theta_k}}\} &= 0, \;\; k = 1, \ldots, K, \\
-\frac{\partial V_k}{\partial t} = \inf_{u \in \mathcal{U}^{\bar{\theta_k}}} &\left(L_k(x,u,t) + \frac{\partial V_k}{\partial x} f_k(x,u) \right), \;\; k = 1, \ldots, K, \\
m(x,t) = \sum_{k=1}^{K} m^{\Theta_k}(x,t) &\left(\int_{\Theta_k} \rho_0(\alpha) d\alpha \right), \\
\bar{x} = \int_{-\infty}^{+\infty} x m(x,t) dx, & \\
q_t = \left| \lambda \int_0^t (\bar{x} - y) dt \right| + k_p, &
\end{aligned} \tag{3.10}$$

where

$$\begin{aligned}
L_k(x,u,t) &= \frac{1}{2} e^{-\delta t} [(x-z)^2 q_t + (x - \bar{\theta_k})^2 q_0 + (u)^2 r], \;\; k = 1, \ldots, K, \\
f_k(x, u_*^{\bar{\theta_k}}) &= -a(x^* - \bar{\theta_k}) + b u_*^{\bar{\theta_k}}, \;\; k = 1, \ldots, K.
\end{aligned}$$

A fixed point must be found for this system, in the sense that starting from a posited family of flows $\left\{m^{\Theta_k}(x,t)\right\}$, one recovers the same $\left\{m^{\Theta_k}(x,t)\right\}$ when solving a family of K advection equations under the associated best response optimal control laws for each k in (3.10).

The decentralized control mechanism is based on the fixed point of the MF system which is calculated offline. At the start of the control horizon, the distribution of initial mean temperature of the population is given so we can compute $\bar{x}_0$. Depending on the status of generation and demand from loads, the central authority will determine the temperature target y and a corresponding parameter z in order to balance the load and generation. For example, when the objective is to mitigate renewable variability, if the renewable generation is predicted to be lower than the peak demand, the central authority will require all devices to decrease their temperatures, and then $z < y < \bar{x}_0$. Based on y, z, and the initial device population temperatures, the central authority can compute the trajectory of q_t at the fixed point, and transmit the information to all devices which can in turn compute their control locally (see Figure 3.1). In practice,

it may be required that the central authority periodically collect samples of home temperatures to update temperature distribution of the population in order to correct any error that may happen during the control horizon (periodic measurements represented by the dotted line in Figure 3.1).

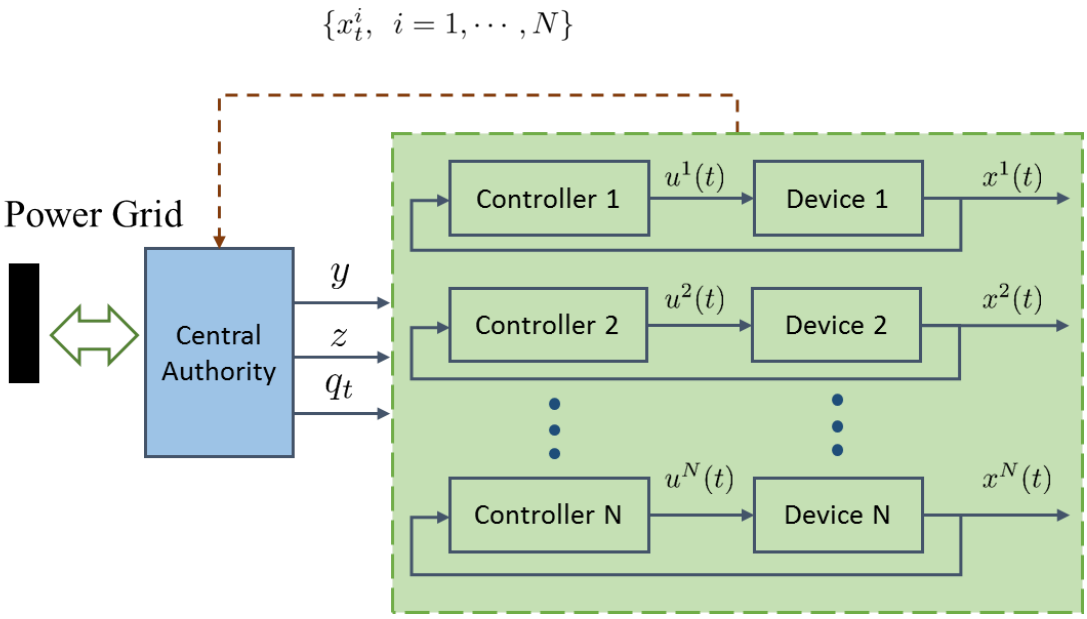

Figure 3.1 Decentralized control mechanism based on fixed point of the MF system

CHAPTER 4 CONTROL POLICY AND COMPUTATION OF THE EQUILIBRIUM SOLUTION

In this chapter, we analyze the structure of the optimal control policy under constrained control inputs. The optimal solution is based on solving the HJB equation associated with the best response of agents given q_t acting on their cost function. We then impose a sufficient condition under which control saturation, if any, occurs only at the start of the control horizon for a finite period of time, then control becomes unsaturated until the end of control horizon. Thus best responses are constrained to be at most single switching strategies. Finally we propose a numerical scheme to compute the single switching control policies for the devices and to solve the Fokker-Planck equation in the MF equation system in (3.10). An iterative algorithm is presented to compute the associated equilibrium solution within the class of at most one switching control strategies.

4.1 Structure of Optimal Control Policies

For a certain agent i, we have the HJB equation defined as in (3.6),

$$-\frac{\partial V_i}{\partial t} = \inf_{u^i \in \mathcal{U}^i} \left(L(x^i, u^i, t) + \frac{\partial V_i}{\partial x^i} f(x^i, u^i) \right),$$

where

$$\begin{aligned} L(x^i, u^i, t) &= \frac{1}{2} e^{-\delta t} [(x^i - z)^2 q_t + (x^i - x_0^i)^2 q_0 + (u^i)^2 r], \\ f(x^i, u^i) &= -a(x^i - x_0^i) + b u^i. \end{aligned}$$

Note from (3.6) that $L(x^i, u^t, t)$ is a time varying function explicitly dependent on t. For the remainder of our analysis, we perform a change of variable to make it implicitly dependent on t. Denote $\tilde{x}^i = e^{-\frac{\delta t}{2}} x^i$, $\tilde{x}_0^i = e^{-\frac{\delta t}{2}} x_0^i$, $\tilde{z} = e^{-\frac{\delta t}{2}} z$,

and $\tilde{u}^i = e^{-\frac{\delta t}{2}} u^i$, we have

$$\begin{aligned} L(x^i, u^i, t) &= \frac{1}{2} e^{-\delta t}[(x^i - z)^2 q_t + (x^i - x_0^i)^2 q_0 + (u^i)^2 r] \\ &= \frac{1}{2}[(\tilde{x}^i - \tilde{z})^2 q_t + (\tilde{x}^i - \tilde{x}_0^i)^2 q_0 + (\tilde{u}^i)^2 r] \\ &= \tilde{L}(\tilde{x}^i, \tilde{u}^i). \end{aligned} \tag{4.1}$$

Accordingly, we have

$$\begin{aligned} V_i(x^i, t) &= \inf_{u^i \in \mathcal{U}^i} J_i(x^i, u^i, t) \\ &= 2 \inf_{u^i \in \mathcal{U}^i} \int_0^\infty L(x^i, u^i, t) dt \\ &= 2 \inf_{\tilde{u}^i \in \tilde{\mathcal{U}}^i} \int_0^\infty \tilde{L}(\tilde{x}^i, \tilde{u}^i) dt \\ &= \inf_{\tilde{u}^i \in \tilde{\mathcal{U}}^i} \tilde{J}_i(\tilde{x}^i, \tilde{u}^i, t) \\ &= \tilde{V}_i(\tilde{x}^i, t). \end{aligned}$$

$$\frac{\partial V_i}{\partial t} = \frac{\partial V_i}{\partial \tilde{V}_i} \frac{\partial \tilde{V}_i}{\partial t} = \frac{\partial \tilde{V}_i}{\partial t}. \tag{4.2}$$

$$\frac{\partial V_i}{\partial x^i} = \frac{\partial V_i}{\partial \tilde{V}_i} \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} \frac{\partial \tilde{x}^i}{\partial x^i} = \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} e^{-\frac{\delta t}{2}}.$$

$$\begin{aligned} \dot{\tilde{x}}^i &= e^{-\frac{\delta t}{2}} (\dot{x}^i - \frac{\delta}{2} x^i) \\ &= e^{-\frac{\delta t}{2}} \left(-a(x^i - x_0^i) + b u^i - \frac{\delta}{2} x^i \right) \\ &= e^{-\frac{\delta t}{2}} \left(-(a + \frac{\delta}{2}) x^i + a x_0^i + b u^i \right) \\ &= -(a + \frac{\delta}{2}) \tilde{x}^i + a \tilde{x}_0^i + b \tilde{u}^i. \end{aligned}$$

$$\begin{aligned}\frac{\partial V_i}{\partial x^i} f(x^i, u^i) &= \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} e^{-\frac{\delta t}{2}} \dot{x}^i \\ &= \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} (\dot{\tilde{x}}^i + \frac{\delta}{2} \tilde{x}^i) \\ &= \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} (-a(\tilde{x}^i - \tilde{x}_0^i) + b\tilde{u}^i) \\ &= \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} f(\tilde{x}^i, \tilde{u}^i).\end{aligned} \tag{4.3}$$

From equations (4.1), (4.2), and (4.3), we can rewrite the HJB equation in (3.6) in terms of the new variables $(\tilde{x}^i, \tilde{u}^i)$, and get

$$-\frac{\partial \tilde{V}_i}{\partial t} = \inf_{\tilde{u}^i \in \tilde{\mathcal{U}}^i} \left(\tilde{L}(\tilde{x}^i, \tilde{u}^i) + \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} f(\tilde{x}^i, \tilde{u}^i) \right). \tag{4.4}$$

The control input constraint we impose to the agent i is in the form $\{u^i \in \mathcal{U} | u_{\min} - u^i_{free} \leq u^i \leq u_{\max} - u^i_{free}\}$, where u^i_{free}, $u_{\min}$ and $u_{\max}$ are defined in Chapter 3. By following the same change of variables, we write the associated input constraint $\{\tilde{u}^i \in \tilde{\mathcal{U}} | \tilde{u}_{\min} - \tilde{u}^i_{free} \leq \tilde{u}^i \leq \tilde{u}_{\max} - \tilde{u}^i_{free}\}$.

In general, an optimal solution $(\tilde{u}^{i*}, \tilde{x}^{i*})$ to the HJB equation in (4.4) can be solved numerically using dynamic programming in a backward direction ($\tilde{u}^{i*}$ is the optimal control and $\tilde{x}^{i*}$ is the corresponding optimal trajectory for agent i) [Kirk, 1970]. Assuming that we are given $(\tilde{u}^{i*}, \tilde{x}^{i*})$ as an optimal solution to (4.4), we have

$$-\frac{\partial}{\partial t} \tilde{V}_i(\tilde{x}^{i*}, t) = \tilde{L}(\tilde{x}^{i*}, \tilde{u}^{i*}) + \frac{\partial}{\partial \tilde{x}^i} \tilde{V}_i(\tilde{x}^{i*}, t) f(\tilde{x}^{i*}, \tilde{u}^{i*}). \tag{4.5}$$

If we differentiate both sides of (4.5) with respect to $\tilde{x}^i$ and use the chain

rule on the right side, we get

$$\begin{aligned} -\frac{\partial}{\partial \tilde{x}^i}\left[\frac{\partial}{\partial t}\tilde{V}_i(\tilde{x}^{i*},t)\right] &= \frac{\partial}{\partial \tilde{x}^i}\tilde{L}(\tilde{x}^{i*},\tilde{u}^{i*}) + \frac{\partial}{\partial \tilde{x}^i}\left[\frac{\partial}{\partial \tilde{x}^i}\tilde{V}_i(\tilde{x}^{i*},t)f(\tilde{x}^{i*},\tilde{u}^{i*})\right], \\ -\frac{\partial^2}{\partial \tilde{x}^i \partial t}\tilde{V}_i(\tilde{x}^{i*},t) &= \frac{\partial}{\partial \tilde{x}^i}\tilde{L}(\tilde{x}^{i*},\tilde{u}^{i*}) + \frac{\partial^2}{\partial \tilde{x}^{i2}}\tilde{V}_i(\tilde{x}^{i*},t)f(\tilde{x}^{i*},\tilde{u}^{i*}) \\ &\quad + \frac{\partial}{\partial \tilde{x}^i}\tilde{V}_i(\tilde{x}^{i*},t)\frac{\partial}{\partial \tilde{x}^i}f(\tilde{x}^{i*},\tilde{u}^{i*}). \end{aligned} \tag{4.6}$$

We now define an *adjoint variable* $\tilde{p}^i(t)$ such that

$$\tilde{p}^i(t) \triangleq \frac{\partial}{\partial \tilde{x}^i}\tilde{V}_i(\tilde{x}^{i*},t). \tag{4.7}$$

By taking the total derivative to $\tilde{p}^i(t)$ in (4.7) with respect to time, we get

$$\begin{aligned} \frac{d\tilde{p}^i}{dt} &= \frac{d}{dt}\left[\frac{\partial}{\partial \tilde{x}^i}\tilde{V}_i(\tilde{x}^{i*},t)\right] \\ &= \frac{\partial}{\partial \tilde{x}^i}\left[\frac{\partial}{\partial x^i}\tilde{V}_i(\tilde{x}^{i*},t)\right]\frac{d\tilde{x}^{i*}}{dt} + \frac{\partial^2}{\partial \tilde{x}^i \partial t}\tilde{V}_i(\tilde{x}^{i*},t) \\ &= \frac{\partial^2}{\partial \tilde{x}^{i2}}\tilde{V}_i(\tilde{x}^{i*},t)f(\tilde{x}^{i*},\tilde{u}^{i*}) + \frac{\partial^2}{\partial \tilde{x}^i \partial t}\tilde{V}_i(\tilde{x}^{i*},t). \end{aligned} \tag{4.8}$$

Note that in (4.8), $\frac{d\tilde{x}^{i*}}{dt} = f(\tilde{x}^{i*},\tilde{u}^{i*})$. Rearranging the terms in (4.6) and using the results in (4.7) and (4.8), we have

$$-\frac{d\tilde{p}^i}{dt} = \frac{\partial}{\partial \tilde{x}^i}\tilde{L}(\tilde{x}^{i*},\tilde{u}^{i*}) + \tilde{p}^i\frac{\partial}{\partial \tilde{x}^i}f(\tilde{x}^{i*},\tilde{u}^{i*}) \tag{4.9}$$

with a terminal condition $\tilde{p}^i(T) = p^i_\infty$, as $T \to \infty$.

Let the Hamiltonian function be defined below, assuming $(\tilde{u}^{i*},\tilde{x}^{i*})$ an optimal solution

$$\tilde{H}(\tilde{p}^i,x^{i*},u^{i*}) = \tilde{L}(\tilde{x}^{i*},\tilde{u}^{i*}) + \tilde{p}^i f(\tilde{x}^{i*},\tilde{u}^{i*}). \tag{4.10}$$

Then for all admissible controls $u^i \in \mathcal{U}^i$, we must have

$$\tilde{H}(\tilde{p}^i, x^{i*}, u^{i*}) \leq \tilde{H}(\tilde{p}^i, x^{i*}, u^i). \tag{4.11}$$

It is therefore shown that given an optimal solution $(\tilde{u}^{i*}, \tilde{x}^{i*})$ to the HJB equation in (4.4), the optimal solution also satisfies equations in (4.9) and (4.11), which are conditions for Pontryagin's Maximum Principle (PMP). As PMP accommodates constraints more easily, we will try to find an optimal control law in an analytic form based on the PMP conditions.

Expanding the PMP inequality in (4.11) by using the Hamiltonian defined in (4.10), we get

$$\begin{aligned}
\tilde{H}(\tilde{p}^i, \tilde{x}^{i*}, \tilde{u}^{i*},) &\leq \tilde{H}(\tilde{p}^i, \tilde{x}^{i*}, \tilde{u}^i), \\
\tilde{L}(\tilde{x}^{i*}, \tilde{u}^{i*}) + \tilde{\tilde{p}}^i f(\tilde{x}^{i*}, \tilde{u}^{i*}) &\leq \tilde{L}(\tilde{x}^{i*}, \tilde{u}^i) + \tilde{p}^i f(\tilde{x}^{i*}, \tilde{u}^{i*}), \\
(\tilde{x}^{i*} - \tilde{z})^2 q_t + (\tilde{x}^{i*} - \tilde{x}_0^i)^2 q_0 + (\tilde{u}^{i*})^2 r + \tilde{p}^i(a(x^{i*} - x_0^i) + b\tilde{u}^{i*}) & \\
\leq (\tilde{x}^{i*} - \tilde{z})^2 q_t + (\tilde{x}^{i*} - \tilde{x}_0^i)^2 q_0 + (\tilde{u}^i)^2 r + \tilde{p}^i(a(x^{i*} - x_0^i) + b\tilde{u}^i), & \\
(\tilde{u}^{i*})^2 r + b\tilde{p}^i \tilde{u}^{i*} &\leq (\tilde{u}^i)^2 r + b\tilde{p}^i \tilde{u}^i.
\end{aligned} \tag{4.12}$$

From the inequality in (4.12), we can derive the following optimal control law for all admissible controls $u^i \in \mathcal{U}^i$ using the earlier change of variables where $\tilde{u}^i = e^{\frac{\delta t}{2}} u^i$, and $\tilde{p}^i = e^{\frac{\delta t}{2}} p^i$:

Optimal Control Law

$$u_*^i = \begin{cases} u_-^i, & h(p^i) < u_-^i \\ h(p^i), & h(p^i) \in [u_-^i, u_+^i] \\ u_+^i, & h(p^i) > u_+^i \end{cases}, \tag{4.13}$$

where for an agent i, $h(p^i) = -br^{-1}p^i$. $u_-^i = u_{\min} - u_{free}^i$ is the lower bound of the constraint $\mathcal{U}^i$, and $u_+^i = u_{\max} - u_{free}^i$ is the upper bound of the constraint. Starting from (4.9) and using the earlier change of variables, we have the following costate equation

$$\dot{p^i} = (a + \delta)p^i - q_t(x^i - z) - q_0(x^i - x_0^i),\ p^i(T) = p_\infty^i. \tag{4.14}$$

Suppose that p^i is in the form of $p^i(t) = \pi^i(t)x^i(t) + s^i(t)$, where $\pi^i(t)$ and $s^i(t)$ are two functions to be determined [Sontag, 2013]. For the case where $u^i = -br^{-1}p^i \in [u^i_-, u^i_+]$, i.e. controls are unconstrained, we have

$$\begin{aligned}
\dot{x}^i &= -a(x^i - x^i_0) + b(-br^{-1}p^i) \\
&= -(a + b^2r^{-1}\pi^i)x^i - b^2r^{-1}s^i + ax^i_0, \\
\dot{p}^i &= \dot{\pi}^i x^i + \pi^i \dot{x}^i + \dot{s}^i, \\
\dot{\pi}^i x^i + \pi^i \dot{x}^i + \dot{s}^i &= (a+\delta)p^i - q_t(x^i - z) - q_0(x^i - x^i_0), \\
\dot{\pi}^i x^i + \dot{s}^i &= -\pi^i\left(-(a + b^2r^{-1}\pi^i)x^i - b^2r^{-1}s^i + ax^i_0\right) \\
&\qquad + (a+\delta)(\pi^i x^i + s^i) - q_t(x^i - z) - q_0(x^i - x^i_0) \\
&= \left((2a+\delta)\pi^i + b^2r^{-1}\pi^{i2} - q_t - q_0\right)x^i \\
&\qquad + \left((a + \delta + b^2r^{-1}\pi^i)s^i - a\pi^i x^i_0 + q_t z + q_0 x^i_0\right).
\end{aligned}$$

Therefore we have

$$\frac{d\pi^i}{dt} = (2a+\delta)\pi^i + b^2r^{-1}\pi^{i2} - q_t - q_0. \tag{4.15}$$

$$\frac{ds^i}{dt} = (a + \delta + b^2r^{-1}\pi^i)s^i - ax^i_0\pi^i + q_t z + q_0 x^i_0. \tag{4.16}$$

Equation (4.15) is a scalar Riccati equation, and we have $\pi^i(t) > 0, \ \forall t > 0$ [Sontag, 2013]. The trajectories of $\pi^i(t)$ and $s^i(t)$ when controls are unconstrained can be solved from equations in (4.15) and (4.16) respectively for a given q_t in a backwards direction subject to some terminal conditions. For an infinite control horizon problem, all terminal conditions equal to zero as there is no terminal cost from the cost function defined in (3.4).

By following a similar approach, we can write the equations for π^i and s^i when controls are constrained. Considering without loss of generality the

temperature decrease case, when controls are constrained, i.e. $u^i = u^i_-$, we have

$$\begin{aligned}
\dot{x}^i &= -a(x^i - x^i_0) + bu^i_-, \\
\dot{p}^i &= \dot{\pi}^i x^i + \pi^i \dot{x}^i + \dot{s}^i, \\
\dot{\pi}^i x^i + \pi^i \dot{x}^i + \dot{s}^i &= (a+\delta)p^i - q_t(x^i - z) - q_0(x^i - x^i_0), \\
\dot{\pi}^i x^i + \dot{s}^i &= -\pi^i \left(-a(x^i - x^i_0) + bu^i_-\right) \\
&\qquad + (a+\delta)(\pi^i x^i + s^i) - q_t(x^i - z) - q_0(x^i - x^i_0) \\
&= \left((2a+\delta)\pi^i - q_t - q_0\right) x^i \\
&\qquad + \left((a+\delta)s^i - a\pi^i x^i_0 + q_t z + q_0 x^i_0 - b\pi^i u^i_-\right).
\end{aligned}$$

Therefore we have

$$\frac{d\pi^i}{dt} = (2a+\delta)\pi^i - q_t - q_0. \tag{4.17}$$

$$\frac{ds^i}{dt} = (a+\delta)s^i - a\pi^i x^i_0 + q_t z + q_0 x^i_0 - b\pi^i u^i_-. \tag{4.18}$$

While based on the optimal control law in (4.13), under a given pressure q_t, the control could go in theory from no saturation to saturation, and then back to no saturation. The pattern could repeat several times before reaching the steady-state. We wish to restrict the set of q_t by imposing conditions on q_t such that saturation, if any, would only be allowed at the start of the control horizon until some time t^i_*. Past t^i_* the optimal control becomes unsaturated for the rest of the control horizon. We now discuss the restrict set of q_t and a sufficient condition for such a single switching control policy.

4.2 Single Switching Control Policy

We still assume without loss of generality the temperature decrease case, where we have $u_{\min} = 0$, $\forall i \in [N]$ such that the total control cannot be a negative value to cool down the space. In addition, we assume $\bar{x}_0 > y > z > x_a$, and $x_0^i > z$, $\forall i \in [N]$. Therefore we have $u_-^i = -u_{free}^i$. We then require $\frac{dp^i}{dt} \leq 0$ for a monotonically decreasing p^i. The reason for a monotonic p^i is that, suppose that at some time t', $-br^{-1}p^i(t') < u_-^i$, then the optimal control is saturated, i.e., $u^{i*} = u_-^i$. When p^i monotonically decreases, the control will become non-saturated at some $t > t'$, and remain so until the end of control horizon. Hence, we eliminate the possibility that the control may go back and forth between saturation and no saturation.

From (4.14), the monotonicity of costate p^i imposes a condition on q_t whereby,

$$q_t \geq \frac{(a+\delta)p^i}{x^i - z} + q_0 \frac{x_0^i - x^i}{x^i - z} \quad \forall i = [N]. \tag{4.19}$$

If the given q_t satisfies (4.19), then we are guaranteed that the best response control policies under such q_t have the single switching properties. Therefore we would like to study under what conditions (4.19) holds. In the following sections, we produce a lower bound q_t^- on q_t and an upper bound for the right hand side of (4.19). In order to do so, we shall assume that δ is small enough that the steady-state mean temperature trajectory $\bar{x}$ asymptotically reaches the target y (if δ is too large, we may arrive at a fixed point solution where $\bar{x}$ settles at the temperature boundary z; however this is not a desired steady-state). At the desired steady-state $\bar{x} = y$, q_t reaches a constant value q_y^∞. It turns out that q_y^∞ is the unique constant weight under which all individuals reach their steady-states, and the mean of the individual steady-states equals to y. Given the initial mean temperature $\bar{x}_0$ of the population, we can compute the q_y^∞, as discussed in the following in details.

4.2.1 Calculation of q_y^∞

Proposition 4.2.1. *At steady-state if $\bar{x}$ settles at y, each agent of the population has an unconstrained control input.*

Proof. Suppose that the control input is constrained for agent i at steady-state, then π^i and s^i follow the ODE's given in (4.17) and (4.18). Given q_y^∞, we can compute the boundary conditions as $T \to \infty$,

$$\begin{aligned}
\dot{\pi}^i &= (2a+\delta)\pi^i - q_t - q_0, \\
\dot{s}^i &= (a+\delta)s^i - a\pi^i x_0^i + q_t z + q_0 x_0^i - b\pi^i u_-^i. \\
&\text{subject to boundary conditions at steady-state} \\
0 &= (2a+\delta)\pi_\infty^i - q_y^\infty - q_0, \\
0 &= (a+\delta)s_\infty^i - a\pi_\infty^i x_0^i + q_y^\infty z + q_0 x_0^i - b\pi_\infty^i u_-^i,
\end{aligned}$$

We have

$$\begin{aligned}
\pi_\infty^i &= \frac{1}{2a+\delta}\left(q_y^\infty + q_0\right), \\
s_\infty^i &= -\frac{1}{a+\delta}\left(-a\pi_\infty^i x_0^i + q_y^\infty z + q_0 x_0^i - b\pi_\infty^i u_-^i\right) \\
&= -\frac{1}{a+\delta}\left(-\pi_\infty^i(ax_0^i + bu_-^i) + q_y^\infty z + q_0 x_0^i\right) \\
&= -\frac{1}{a+\delta}\left(-\pi_\infty^i(ax_0^i - bu_{free}^i) + q_y^\infty z + q_0 x_0^i\right) \\
&= -\frac{1}{a+\delta}\left(-\pi_\infty^i(ax_0^i - bb^{-1}a(x_0^i - x_a)) + q_y^\infty z + q_0 x_0^i\right) \\
&= -\frac{1}{a+\delta}\left(-a\pi_\infty^i x_a + q_y^\infty z + q_0 x_0^i\right).
\end{aligned}$$

When controls are constrained, the inequality $-br^{-1}p_\infty^i < u_-^i$ must hold at steady-state, where

$$
\begin{aligned}
p^i_\infty &= \pi^i_\infty x^i_\infty + s^i_\infty \\
&= \pi^i_\infty x_a + s^i_\infty \\
&= \pi^i_\infty x_a - \frac{1}{a+\delta}\left(-a\pi^i_\infty x_a + q^\infty_y z + q_0 x^i_0\right) \\
&= \frac{1}{a+\delta}\left((2a+\delta)\pi^i_\infty x_a - q^\infty_y z - q_0 x^i_0\right) \\
&= \frac{1}{a+\delta}\left((2a+\delta)\left[\frac{1}{2a+\delta}(q^\infty_y + q_0)\right]x_a - q^\infty_y z - q_0 x^i_0\right) \\
&= -\frac{1}{a+\delta}\left(q^\infty_y(z-x_a) + q_0(x^i_0 - x_a)\right) \\
&< 0
\end{aligned}
$$

given that we have $z - x_a > 0$ and $x^i_0 - x_a > 0$.

Consequently we have

$$-br^{-1}p^i_\infty = br^{-1}\left[\frac{1}{a+\delta}\left(q^\infty_y(z-x_a) + q_0(x^i_0 - x_a)\right)\right] > 0 > u^i_-.$$

However, the above is in contradiction with the inequality condition that when controls are constrained we should have $-br^{-1}p^i_\infty < u^i_- < 0$. Recall that $u^i_- = -u^i_{free} = -b^{-1}a(x^i_0 - x_a) < 0$ in the temperature decrease case. Hence we can conclude that at steady-state, control inputs cannot be constrained for any agent of the population. □

By Proposition 4.2.1, we can conclude that in the constrained case, all controls become unconstrained at steady-state. Therefore, all quantities at steady-states should be identical in both the constrained and unconstrained cases.

Proposition 4.2.2. *Let* $c_1 = \frac{a(a+\delta)r + b^2 q_0}{b^2}$*, the unique* q^∞_y *can be computed by*

$$q^\infty_y = c_1 \frac{\bar{x}_0 - y}{y - z}.$$

Proof. Suppose that at steady-state the mean temperature of the population

$\bar{x}$ reaches the target y, while control input for any device of the population should be unconstrained from Proposition 4.2.1. Therefore we have optimal control $u_*^i = -br^{-1}p^i$, where π^i and s^i are described by (4.15) and (4.16). By substituting the optimal control into the process equation, we get,

$$\frac{dx^i}{dt} = -(a + b^2r^{-1}\pi^i)x^i - b^2r^{-1}s^i + ax_0^i. \tag{4.20}$$

At steady-state, we should have $\frac{d\pi^i}{dt} = 0$, $\frac{ds^i}{dt} = 0$, and $\frac{dx^i}{dt} = 0$. Then we obtain the following from (4.15)-(4.20),

$$-(2a + \delta)\pi_\infty^i - b^2r^{-1}\pi_\infty^{i2} + q_y^\infty + q_0 = 0, \tag{4.21}$$

$$-(a + \delta + b^2r^{-1}\pi_\infty^i)s_\infty^i + ax_0^i\pi_\infty^i - q_y^\infty z - q_0x_0^i = 0, \tag{4.22}$$

$$-(a + b^2r^{-1}\pi_\infty^i)x_\infty^i - b^2r^{-1}s_\infty^i + ax_0^i = 0, \tag{4.23}$$

where π_∞^i, s_∞^i, and x_∞^i denote values at steady-state. To characterize the q_y^∞ in terms of x_∞^i, we first multiply (4.21) by x_∞^i and get

$$-(2a + \delta)\pi_\infty^i x_\infty^i - b^2r^{-1}\pi_\infty^{i2}x_\infty^i + q_y^\infty x_\infty^i + q_0x_\infty^i = 0. \tag{4.24}$$

Then by adding (4.24) to (4.22) we get

$$\begin{aligned}-(a + \delta + b^2r^{-1}\pi_\infty^i)s_\infty^i + ax_0^i\pi_\infty^i - (2a + \delta)\pi_\infty^i x_\infty^i - b^2r^{-1}\pi_\infty^{i2}x_\infty^i \\ +q_y^\infty(x_\infty^i - z) + q_0(x_\infty^i - x_0^i) = 0.\end{aligned} \tag{4.25}$$

From (4.23), we get

$$s_\infty^i = \frac{-(a + b^2r^{-1}\pi_\infty^i)x_\infty^i + ax_0^i}{b^2r^{-1}}. \tag{4.26}$$

Substituting the result of (4.26) into (4.25), we get

$$\begin{aligned}-(a + \delta + b^2r^{-1}\pi_\infty^i)\left[-(a + b^2r^{-1}\pi_\infty^i)x_\infty^i + ax_0^i\right] - (2a + \delta)b^2r^{-1}\pi_\infty^i x_\infty^i \\ +ab^2r^{-1}x_0^i\pi_\infty^i - (b^2r^{-1})^2\pi_\infty^{i2}x_\infty^i + b^2r^{-1}\left[q_y^\infty(x_\infty^i - z) + q_0(x_\infty^i - x_0^i)\right] = 0.\end{aligned} \tag{4.27}$$

Expanding the terms in (4.27), we get,

$$\begin{aligned} a(a+\delta)(x^i_\infty - x^i_0) + b^2 r^{-1}\left[q^\infty_y(x^i_\infty - z) + q_0(x^i_\infty - x^i_0)\right] + (2a+\delta)b^2 r^{-1}\pi^i_\infty x^i_\infty \\ +(b^2 r^{-1})^2 \pi^{i2}_\infty x^i_\infty - ab^2 r^{-1} x^i_0 \pi^i_\infty + ab^2 r^{-1} x^i_0 \pi^i_\infty - (2a+\delta)b^2 r^{-1}\pi^i_\infty x^i_\infty \\ -(b^2 r^{-1})^2 \pi^{i2}_\infty x^i_\infty = 0. \end{aligned} \tag{4.28}$$

Only the first two terms remain while the latter terms in (4.28) cancel each other out, leading to the following expression for q^∞_y

$$q^\infty_y = c_1 \frac{x^i_0 - x^i_\infty}{x^i_\infty - z}, \tag{4.29}$$

where

$$c_1 = \frac{a(a+\delta)r + b^2 q_0}{b^2}.$$

Multiplying by $(x^i_\infty - z)$ on both sides of (4.29), we get

$$(x^i_\infty - z)q^\infty_y = c_1(x^i_0 - x^i_\infty).$$

By taking expected values on both sides of the above equation given $\mathbb{E}\{x^i_\infty\} = \bar{x}_\infty = y$ at steady-state and q^∞_t is a constant, we get

$$q^\infty_t = c_1 \frac{\bar{x}_0 - y}{y - z}. \tag{4.30}$$

Hence we prove Proposition 4.2.2. □

In the following analysis, we take the computed q^∞_t as the upper bound of q_t when searching for the fixed point. Hence, it is expected that $\bar{x}(t) \leq y, \ \forall t$, and there should be no oscillation around y. In the next section, we produce a lower bound for q_t and an upper bound for the RHS of inequality in (4.19), from which we define the sufficient condition for the single switch control policy. Under a given q_t, we define

$$\Psi(x^i) = \frac{(a+\delta)p^i}{x^i - z} + q_0 \frac{x^i_0 - x^i}{x^i - z}. \tag{4.31}$$

Note that as $p^i = \pi^i x^i + s^i$, we define $\Psi(x^i)$ a function of x^i only.

4.2.2 Upper Bound for $\Psi(x^i)$ and Lower Bound for q_t

Suppose that we apply a constant $q_t(t) = q_y^\infty$, $\forall t$ to (4.14), and we denote $x_-^i(t)$ and $p_-^i(t)$ the resulting state and costate trajectories of any agent. We have

$$\frac{dp_-^i}{dt} = \alpha p_-^i - (x_-^i - z)q_y^\infty - (x_-^i - x_0^i)q_0, \tag{4.32}$$

where $\alpha = a + \delta$. By differentiating both sides of (4.32) with respect to time t, we get

$$\frac{d^2p_-^i}{dt^2} = \alpha\frac{dp_-^i}{dt} - (q_y^\infty + q_0)\dot{x}_-^i. \tag{4.33}$$

For $t \in [0, T]$ where $T \to \infty$, the solution to (4.33) can be expressed as

$$\begin{aligned}\frac{dp_-^i}{dt} &= e^{\alpha(t-T)}\dot{p}_-^i(T) + \int_T^t e^{\alpha(t-\tau)}[-(q_y^\infty + q_0)]\dot{x}_-^i(\tau)d\tau \\ &= -(q_y^\infty + q_0)e^{\alpha t}\int_T^t e^{-\alpha\tau}\dot{x}_-^i(\tau)d\tau \\ &= (q_y^\infty + q_0)e^{\alpha t}\int_t^T e^{-\alpha\tau}\dot{x}_-^i(\tau)d\tau.\end{aligned} \tag{4.34}$$

Note that $\dot{p}_-^i(T) = \dot{\pi}^i(T)x^i(T) + \pi^i(T)\dot{x}^i(T) + \dot{s}^i(T)$, as $\dot{\pi}^i$, $\dot{x}^i$, and $\dot{s}^i$=0 at $T \to \infty$, we have $\dot{p}_-^i(T) = 0$.

Under q_y^∞, x_-^i is monotonically decaying (see (4.46) below for expression of x_-^i), hence $\dot{x}_-^i \leq 0$, $\forall t$, and we verify $\frac{dp_-^i}{dt} \leq 0$, $\forall t$. This indicates that when $q_t(t) = q_y^\infty$ is applied, (4.19) is satisfied.

From the produced state trajectory $x_-^i(t)$ and costate trajectory $p_-^i(t)$, we produce an upper bound

$$\Psi(x_-^i) = \frac{(a+\delta)p_-^i}{x_-^i - z} + q_0\frac{x_0^i - x_-^i}{x_-^i - z},$$

such that

$$\Psi(x_-^i) \geq \Psi(x^i), \tag{4.35}$$

where x^i is the state trajectory produced under any $q_t \leq q_y^\infty,\ \forall t$, where q_t is defined by (3.5). To see this, when we apply the maximally possible pressure q_y^∞ to ask each device to decrease the temperature, we are applying the most aggressive control inputs, hence we must have $z < x_-^i \leq x^i$ and $p_-^i \geq p^i > 0$. For the right hand side of inequality in (4.35), $\Psi(x_-^i)$ is obtained when x^i and p^i are replaced in $\Psi(x^i)$ by x_-^i and p_-^i respectively, and we are increasing the numerator and decreasing the denominator. Hence we must have $\Psi(x_-^i) \geq \Psi(x^i)$.

By applying the same constant q_y^∞ to $\bar{x}_0$, we can produce a trajectory $\bar{x}_-(t)$ below $\bar{x}(t)$, based on which we can compute a lower bound q_t^- for q_t such that,

$$\begin{aligned} q_t^- &= \lambda \left| \int_0^t (\bar{x}_-(\tau) - y) d\tau \right| \\ &\leq \lambda \left| \int_0^t (\bar{x}(\tau) - y) d\tau \right| \\ &\leq \lambda \left| \int_0^t (\bar{x}(\tau) - y) d\tau \right| + k_p(t) \\ &= q_t. \end{aligned} \tag{4.36}$$

Note that as we consider $q_t \leq q_y^\infty,\ \forall t$, we do not expect any oscillation of $\bar{x}$ around y before reaching steady-state. Further, given that each x_-^i under q_y^∞ is monotonic towards its steady-state, $\bar{x}_-$ should not oscillate around y either.

4.2.3 Sufficient Condition for the Single Switching Control Policy

By adding a proper time varying positive term $k_c(t)$ to q_t^-, we guarantee that the inequality

$$q_t^- + k_c \geq \Psi(x_-^i)\ \ \forall i = [N]$$

holds for all t. As we have shown in the previous section that when q_y^∞ is applied, $\frac{dp_-^i}{dt} \leq 0$, $\forall i \in [N]$, $\forall t$, we must have $q_y^\infty \geq \Psi(x_-^i),\ \forall t$. This indicates that such $k_c(t)$ term always exists to guarantee that the inequality above holds, i.e. $\Psi(x_-^i) - q_t^- \leq k_c < q_y^\infty - q_t^-,\ \forall t$. Therefore by combining (4.19), (4.35), and (4.36) we can write a sufficient condition under which $p^i(t)$ is monotonically

decreasing for the temperature decrease case.

$$\begin{aligned} q_t + k_c \geq q_t - k_p + k_c &\geq q_t^- + k_c, \\ &\geq \Psi(x_-^i) \geq \Psi(x^i) \ \ \forall i = [N]. \end{aligned} \tag{4.37}$$

Note that if we desire $\bar{x}$ to monotonically reach the target y (no oscillations around y), then the term $(q_t + k_c)$ would be restricted to no larger than q_y^∞ during the entire control horizon. Therefore, $k_c(t)$ should satisfy $\Psi(x_-^i) - q_t^- \leq k_c < q_y^\infty - q_t, \ \forall t$. In (4.37), $\Psi(x_-^i)$ can be computed by explicitly expressing the trajectories of x_-^i and p_-^i, which is discussed in the following sections by adopting the single switching control policy.

Calculation of x_-^i

Suppose that we apply a constant q_y^∞ to the system, we can then write the following equations for $t \in [t_*^i, T]$, where t_*^i is the time when the control input becomes unconstrained. For simplicity and easiness to read, we omit the subscript "_" to all variables in the following analysis of this subsection.

When controls are unconstrained, π^i and s^i follow the equations in (4.15) and (4.16), subject to terminal conditions π_∞^i and s_∞^i respectively.

If we work with an infinite control horizon, i.e. $T \to \infty$, (4.15) becomes an algebraic Riccati equation, hence we can solve for a constant and positive solution π_∞. We have

$$\pi_\infty = \frac{-(2a+\delta) + \sqrt{(2a+\delta)^2 + 4b^2 r^{-1}(q_t^\infty + q_0)}}{2b^2 r^{-1}}. \tag{4.38}$$

Note that π_∞ is independent of agent i. By substituting π_∞ into (4.16), we can solve for $s^i(t)$ from the resulting equation:

$$\begin{aligned} \frac{ds^i}{dt} &= \beta s^i + \gamma^i, \\ \beta &= (a + \delta + b^2 r^{-1} \pi_\infty), \\ \gamma^i &= -a x_0^i \pi_\infty + q_y^\infty z + q_0 x_0^i. \end{aligned} \tag{4.39}$$

The boundary condition s_∞^i of (4.39) is obtained by setting $\frac{ds^i}{dt} = 0$ where steady-state is reached, and we get $s_\infty^i = -\frac{\gamma^i}{\beta}$. The solution to (4.39) subject to s_∞^i is therefore expressed by

$$\begin{aligned} s^i(t) &= e^{\beta(t-T)} s_\infty^i + \int_T^t e^{\beta(t-\tau)} \gamma^i d\tau \\ &= e^{\beta(t-T)} s_\infty^i + \gamma^i e^{\beta t} \int_T^t e^{-\beta\tau} d\tau \\ &= e^{\beta(t-T)} s_\infty^i + e^{\beta t} (-\frac{\gamma^i}{\beta}) \left[e^{-\beta\tau} \Big|_T^t \right] \\ &= e^{\beta(t-T)} s_\infty^i + s_\infty^i (1 - e^{\beta(t-T)}) \\ &= s_\infty^i. \end{aligned} \tag{4.40}$$

The resulting s^i is therefore time invariant but dependent on initial temperature x_0^i of each agent. We can therefore have the costate p^i for an agent i written as

$$p^i(t) = \pi_\infty x^i(t) + s_\infty^i, \;\; t \in [t_*^i, \infty)$$

when control is not saturated.

We already know that when under unsaturated control, i.e. $u^i \geq u_-^i$, the optimal control has the form of $u_*^i = -b r^{-1} p^i$. This indicates that if, at the beginning of the control, the costate $p(0) = \pi_\infty x_0^i + s_\infty^i$ is no greater than some $p_{\max}^i = -b^{-1} r u_-^i$, then during the entire control horizon, the control will never be saturated as $p^i(t)$ is monotonically decreasing when a constant q_y^∞ is applied. In this case, one can rewrite the process equation by substituting the resulting

optimal control and get

$$\begin{aligned}\dot{x}^i &= -a(x^i - x_0^i) - b^2 r^{-1}(\pi_\infty x^i + s_\infty^i) \\ &= (-a - b^2 r^{-1}\pi_\infty)x^i + (ax_0^i - b^2 r^{-1} s_\infty^i) \\ &= \kappa_1 x^i + \kappa_2^i, \;\; t \in [0, \infty).\end{aligned} \tag{4.41}$$

One can solve for x^i in (4.41) subject to initial condition $x^i(0) = x_0^i$, and get

$$\begin{aligned}x^i &= e^{\kappa_1 t} x_0^i + \int_0^t e^{\kappa_1 (t-\tau)} \kappa_2^i d\tau, \\ &= e^{\kappa_1 t} x_0^i + e^{\kappa_1 t} \kappa_2^i \int_0^t e^{-\kappa_1 \tau} d\tau \\ &= e^{\kappa_1 t} x_0^i - \frac{\kappa_2^i}{\kappa_1} e^{\kappa_1 t} e^{-\kappa_1 \tau} \Big|_0^t \\ &= e^{\kappa_1 t} x_0^i - \frac{\kappa_2^i}{\kappa_1} e^{\kappa_1 t}(e^{-\kappa_1 t} - 1) \\ &= (x_0^i + \frac{\kappa_2^i}{\kappa_1}) e^{\kappa_1 t} - \frac{\kappa_2^i}{\kappa_1} \;\; \forall t \in [0, \infty).\end{aligned} \tag{4.42}$$

On the other hand, if at $t = 0$, $p(0)^i > p_{\max}^i$, then the control will be saturated in the beginning, until some time t_*^i where $p^i(t_*^i)$ reduces to $p_{\max}^i$. Then the control will become unsaturated until the end of the control horizon. We can compute the state trajectory from the process equation by using the saturated control $u^i = u_-^i = -u_{free}^i = -ab^{-1}(x_0^i - x_a)$ for $t \in [0, t_*^i]$.

$$\begin{aligned}\dot{x}_{con}^i &= -a(x_{con}^i - x_0^i) + bu_-^i \\ &= -a(x_{con}^i - x_0^i) - b\frac{a}{b}(x_0^i - x_a) \\ &= -ax_{con}^i + ax_a.\end{aligned} \tag{4.43}$$

We use a subscript *con* to specify that the state trajectory is under constrained control. Using the initial condition $x(0) = x_0^i$, the solution to (4.43)

is,

$$\begin{aligned} x^i_{con} &= e^{-at}x^i_0 + \int_0^t e^{-a(t-\tau)}(ax_a)d\tau \\ &= e^{-at}x^i_0 + ax_a e^{-at}\int_0^t e^{a\tau}d\tau \\ &= e^{-at}x^i_0 + x_a e^{-at}(e^{at}-1) \\ &= (x^i_0 - x_a)e^{-at} + x_a,\ t \in [0, t^i_*]. \end{aligned} \tag{4.44}$$

For $t \in (t^i_*, T]$, $p^i(t) < p^i_{\max}$, and the control will be unconstrained. The unconstrained state trajectory x^i_{unc} can be explicitly determined by solving (4.41) and using $x^i_{con}(t^i_*)$ as the initial condition. We get

$$x^i_{unc} = \left(x^i_{con}(t^i_*) + \frac{\kappa^i_2}{\kappa_1}\right)e^{\kappa_1(t-t^i_*)} - \frac{\kappa^i_2}{\kappa_1},\ \ t \in (t^i_*, \infty). \tag{4.45}$$

Therefore the complete state trajectory $x^i_-(t)$ under q^∞_y for $t \in [0,\ \infty)$ when the control is constrained at $t = 0$ is defined by

$$x^i_-(t) = \begin{cases} (x^i_0 - x_a)e^{-at} + x_a, & \text{for } t \in [0, t^i_*] \\ \left(x^i_-(t^i_*) + \frac{\kappa^i_2}{\kappa_1}\right)e^{\kappa_1(t-t^i_*)} - \frac{\kappa^i_2}{\kappa_1}, & \text{for } t \in (t^i_*, \infty), \end{cases} \tag{4.46}$$

where $\kappa_1 = (-a - b^2r^{-1}\pi_\infty)$ and $\kappa^i_2 = (ax^i_0 - b^2r^{-1}s^i_\infty)$. We observe from equations (4.42) and (4.46) that, for an initial condition x^i_0, whether the control is constrained or not in the beginning, the same steady-state value is reached, which is $-\frac{\kappa^i_2}{\kappa_1}$. This result is consistent with Proposition 4.2.1. We can also use (4.46) to describe trajectory of x^i_- when the control is not saturated in the beginning, by simply setting $t^i_* = 0$.

From the equality $p^i(t^i_*) = p^i_{\max}$ we can explicitly compute t^i_*:

$$\begin{aligned}
\pi_\infty \left[(x^i_0 - x_a)e^{-at^i_*} + x_a\right] + s^i_\infty &= -b^{-1}ru^i_-, \\
(x^i_0 - x_a)e^{-at^i_*} &= \frac{1}{\pi_\infty}\left(ab^{-2}r(x^i_0 - x_a) - s^i_\infty\right) - x_a, \\
e^{-at^i_*} &= \frac{\frac{1}{\pi_\infty}\left(ab^{-2}r(x^i_0 - x_a) - s^i_\infty\right) - x_a}{x^i_0 - x_a}, \\
t^i_* &= -\frac{1}{a}\ln A^i,
\end{aligned} \tag{4.47}$$

where $A^i = \frac{\frac{1}{\pi_\infty}\left(ab^{-2}r(x^i_0 - x_a) - s^i_\infty\right) - x_a}{x^i_0 - x_a}$.

Calculation of p^i_-

When control input is unconstrained for $t \in (t^i_*, \infty]$, $p^i_-(t)$ can be expressed as $p^i_-(t) = \pi_\infty x^i_-(t) + s^i_\infty$.

Under constrained control input for $t \in [0, t^i_*]$, trajectories of $\pi^i(t)$ and $s^i(t)$ are described in (4.17) and (4.18) , subject to boundary conditions π_∞ and s^i_∞ respectively. We can then solve for $\pi^i(t)$ and $s^i(t)$ in a backwards direction for $t \in [0, t^i_*]$. Hence, the complete trajectory of p^i_- for $t \in [0, T]$ can be computed in 2 segments.

4.2.4 Monotone Properties of $\Psi(\bullet)$

Having explicitly computed x^i_- and p^i_-, we would like to further analyze the sufficient condition defined in (4.37).

Proposition 4.2.3. *For any two state trajectories x_1, $x_2 \in \{x^i,\ i \in [N]\}$, if $x_{1_0} > x_{2_0}$, then $x_1 > x_2$, $\forall t$.*

Proof. At any time t, $x_1(t)$ and $x_2(t)$ have the same pressure coefficient $q_t(t)$ to move towards the target y. Suppose that at certain time t', $x_1(t') = x_2(t')$, then they will experience the same pressure $q_t(t')(x_1(t') - z)^2 = q_t(t')(x_2(t') - z)^2$

to decrease their temperatures. However, as $x_{1_0} > x_{2_0}$, x_1 will have a larger pressure to move towards its initial condition x_{1_0} than x_2 due to the fact that $q_0(x_1(t') - x_{1_0})^2 > q_0(x_2(t') - x_{2_0})^2$. Therefore, at any time when x_1 tends to come close to x_2, it will experience a larger "pull force" towards x_{1_0}, so it will never come across x_2. Hence $x_1 > x_2$, $\forall t$. □

We state the following assumption.

A1: For the chosen parameters y, z, a, δ, b, r, and q_0, let the following inequality hold:

$$\frac{\bar{x}_0 - y}{y - z} \leq \left(\frac{a}{a+\delta} q_0\right) \Big/ \left(\frac{a(a+\delta)r}{b^2} + q_0\right).$$

Proposition 4.2.4. *Under **A1**, there exists an upper bound for the family of trajectories* $\left\{\Psi(x_-^i),\ \forall i \in [N]\right\}$ *under* q_y^∞*, which is*

$$\max_{i\in[N]} \left\{\Psi(x_-^i)\right\} = \Psi(x_-^{\min}) = \frac{(a+\delta)p_-^{\min}}{x_-^{\min} - z} + q_0 \frac{x_0^{\min} - x_-^{\min}}{x_-^{\min} - z} \ \ \forall t,$$

where $x_-^{\min}$ *denotes the state trajectory starting from the minimum initial temperature* $x_0^{\min}$*, and* $p_-^{\min}$ *is the corresponding costate trajectory.*

Proof. Assume we are given x_{1_0} and x_{2_0}, $x_{1_0} > x_{2_0}$, by Proposition 4.2.3, we have $x_{1-} > x_{2-}$. We want to show

$$\frac{(a+\delta)p_{1-}}{x_{1-} - z} + q_0 \frac{x_{1_0} - x_{1-}}{x_{1-} - z} < \frac{(a+\delta)p_{2-}}{x_{2-} - z} + q_0 \frac{x_{2_0} - x_{2-}}{x_{2-} - z}. \tag{4.48}$$

From the costate equation,

$$\dot{p}^i = (a+\delta)p^i - q_t(x^i - z) - q_0(x^i - x_0^i),\ p^i(T) = p_\infty^i,$$

(4.48) is equivalent to

$$\frac{\dot{p}_{1-}}{x_{1-} - z} < \frac{\dot{p}_{2-}}{x_{2-} - z} \tag{4.49}$$

under a constant pressure q_y^∞. Defining the following function

$$\Gamma(x_-^i) = \frac{\dot{p}_-^i}{x_-^i - z},$$

if we can show $\frac{d}{dx_-^i}\Gamma(x_-^i) < 0$, then for $x_{1-} > x_{2-}$, we must have $\Gamma(x_{1-}) < \Gamma(x_{2-})$, hence Proposition 4.2.4 is proved.

We first consider the case where x_{1-} and x_{2-} are both unconstrained.

$$\begin{aligned}
\Gamma(x_-) &= \frac{\dot{p}_-^i}{x_-^i - z} \\
&= \frac{\frac{d}{dt}(\pi_\infty x_-^i + s_\infty^i)}{x_-^i - z} \\
&= \frac{\pi_\infty \dot{x}_-^i}{x_-^i - z} \\
&= \frac{\pi_\infty[-(a + b^2 r^{-1}\pi_\infty)x_-^i + a x_0^i - b^2 r^{-1} s_\infty^i]}{x_-^i - z}.
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dx_-^i}\Gamma(x_-^i) &= \frac{\pi_\infty}{x_-^i - z}\left[-(a + b^2 r^{-1}\pi_\infty)\right] \\
&\quad - \frac{\pi_\infty}{(x_-^i - z)^2}\left[-(a + b^2 r^{-1}\pi_\infty)x_-^i + a x_0^i - b^2 r^{-1} s_\infty^i\right] \\
&= \frac{\pi_\infty}{(x_-^i - z)^2}\left[-(a + b^2 r^{-1}\pi_\infty)(x_-^i - z)\right] \\
&\quad + \frac{\pi_\infty}{(x_-^i - z)^2}\left[(a + b^2 r^{-1}\pi_\infty)x_-^i - a x_0^i + b^2 r^{-1} s_\infty^i\right] \\
&= \frac{\pi_\infty}{(x_-^i - z)^2}\left[a(z - x_0^i) + b^2 r^{-1}(\pi_\infty z + s_\infty^i)\right] \\
&= \frac{\pi_\infty}{(x_-^i - z)^2}\left[a(z - x_0^i) + b^2 r^{-1}(\pi_\infty z + \frac{-(a + b^2 r^{-1}\pi_\infty^i)x_{\infty-}^i + a x_0^i}{b^2 r^{-1}})\right] \\
&= \frac{\pi_\infty}{(x_-^i - z)^2}\left[a(z - x_0^i) + b^2 r^{-1}\pi_\infty z - (a + b^2 r^{-1}\pi_\infty^i)x_{\infty-}^i + a x_0^i\right] \\
&= \frac{\pi_\infty}{(x_-^i - z)^2}\left[(a + b^2 r^{-1}\pi_\infty)(z - x_{\infty-}^i)\right] \\
&< 0 \ \ \forall x_-^i.
\end{aligned} \tag{4.50}$$

In the case where x_{1-} and x_{2-} are both constrained, we have

$$\begin{aligned}
\Gamma(x_-) &= \frac{\dot{p}^i_-}{x^i_- - z} \\
&= \frac{(a+\delta)p^i_- - q_0(x^i_- - x^i_0)}{x^i_- - z} - q_y^\infty, \\
\frac{d}{dx^i_-}\Gamma(x^i_-) &= \frac{1}{x^i_- - z}\left[(a+\delta)\frac{\partial p^i_-}{dx^i_-} - q_0\right] - \frac{1}{(x^i_- - z)^2}\left[(a+\delta)p^i_- - q_0(x^i_- - x^i_0)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[((a+\delta)\frac{\partial p^i_-}{dx^i_-} - q_0)(x^i_- - z) - (a+\delta)p^i_- + q_0(x^i_- - x^i_0)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[(a+\delta)\frac{\partial p^i_-}{dx^i_-}(x^i_- - z) - (a+\delta)p^i_- + q_0(z - x^i_0)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[(a+\delta)(\frac{\partial p^i_-}{dx^i_-}(x^i_- - z) - p^i_-) + q_0(z - x^i_0)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[(a+\delta)(\pi^i x^i_- - \pi^i z - \pi^i x^i_- - s^i) + q_0(z - x^i_0)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[-(a+\delta)(\pi^i z + s^i) + q_0(z - x^i_0)\right].
\end{aligned} \tag{4.51}$$

Let $z = x^i_- - \Delta x^i$, where $\Delta x^i \in (0, x^i_- - x_a)$. Substituting it into (4.51) we get

$$\begin{aligned}
\frac{d}{dx^i_-}\Gamma(x^i_-, t) &= \frac{1}{(x^i_- - z)^2}\left[-(a+\delta)(\pi^i x^i_- + s^i - \pi^i \Delta x^i) + q_0(x^i_- - x^i_0 - \Delta x^i)\right] \\
&= \frac{1}{(x^i_- - z)^2}\left[-(a+\delta)p^i_- + q_0(x^i_- - x^i_0) + ((a+\delta)\pi^i - q_0)\Delta x^i\right].
\end{aligned} \tag{4.52}$$

By imposing the condition $(a+\delta)\pi^i - q_0 \leq 0$, we have $\frac{d}{dx^i_-}\Gamma(x^i_-) < 0$ for x^i_- under constrained control inputs. For this condition to hold, we impose $q_0 \geq (a+\delta)\sup\{\pi^i\}$. The supreme of π^i happens at $t = 0$ when $t^i_* = T$, which is independent of agent i and when the control remains constrained just before the end of the control horizon. From the solution to (4.17), we can compute

$\sup\{\pi\}$ by having $t_* = T$ which goes to ∞.

$$\begin{aligned}
\pi &= e^{(2a+\delta)(t-t_*)}\pi_\infty - \int_{t_*}^{t} e^{(2a+\delta)(t-\tau)}(q_y^\infty + q_0)d\tau \\
&= e^{(2a+\delta)(t-t_*)}\pi_\infty - (q_y^\infty + q_0)e^{(2a+\delta)t}\int_{t_*}^{t} e^{-(2a+\delta)\tau}d\tau \\
&= e^{(2a+\delta)(t-t_*)}\pi_\infty + \frac{q_y^\infty + q_0}{2a+\delta}e^{(2a+\delta)t}e^{-(2a+\delta)\tau}\Big|_{t_*}^{t} \\
&= \left(\pi_\infty - \frac{q_y^\infty + q_0}{2a+\delta}\right)e^{(2a+\delta)(t-t_*)} + \frac{q_y^\infty + q_0}{2a+\delta}, \\
\sup\{\pi\} &= \frac{q_y^\infty + q_0}{2a+\delta}, \text{ as } t_* \to \infty.
\end{aligned}$$

Note that $\sup\{x\}$ happens when $t = 0$. Therefore, we should have

$$\begin{aligned}
q_0 &\geq (a+\delta)\sup\{\pi\} \\
&= \frac{a+\delta}{2a+\delta}(q_y^\infty + q_0) \\
&= \frac{a+\delta}{2a+\delta}(q_y^\infty + q_0), \\
(2a+\delta)q_0 &\geq (a+\delta)\left[\left(\frac{a(a+\delta)r}{b^2} + q_0\right)\frac{\bar{x}_0 - y}{y-z} + q_0\right], \\
\frac{a}{a+\delta}q_0 &\geq \left(\frac{a(a+\delta)r}{b^2} + q_0\right)\frac{\bar{x}_0 - y}{y-z}, \\
\frac{\bar{x}_0 - y}{y-z} &\leq \left(\frac{a}{a+\delta}q_0\right) \Big/ \left(\frac{a(a+\delta)r}{b^2} + q_0\right). \qquad (4.53)
\end{aligned}$$

As long as the parameters are properly chosen to satisfy the condition in (4.53), we have $\frac{d}{dx_-^i}\Gamma(x_-^i) < 0$ for x_-^i under constrained control inputs.

Therefore we have shown $\frac{d}{dx_-^i}\Gamma(x_-^i) < 0$ in either constrained or unconstrained case. Hence we have $\max_{i\in[N]}\left\{\Psi(x_-^i)\right\} = \Psi(x_-^{\min})$ and we complete proof of Proposition 4.2.4. □

By Proposition 4.2.4, if we choose $\Psi(x_-^{\min}) - q_t^- \leq k_c < q_y^\infty - q_t^-$ we can guarantee that (4.37) holds for single switching control policy, where $\Psi(x_-^{\min})$ and q_t^- can be explicitly computed.

Proposition 4.2.5. *Under **A1**, there exists an upper bound for the family of trajectories* $\left\{\Psi(x^i), \forall i \in [N]\right\}$ *under* $q_t(t) \leq q_t^{\infty}$*, which is*

$$\max_{i\in[N]}\left\{\frac{(a+\delta)p^i}{x^i-z}+q_0\frac{x_0^i-x^i}{x^i-z}\right\}=\frac{(a+\delta)p^{\min}}{x^{\min}-z}+q_0\frac{x_0^{\min}-x^{\min}}{x^{\min}-z}.$$

Proof. We follow the same approach to prove Proposition 4.2.4 by defining a function $\Gamma(x^i)=\frac{p^i}{x^i-z}$, and show $\frac{d}{dx^i}\Gamma(x^i)<0$ in both constrained and unconstrained cases. We can show that in both cases,

$$\frac{d}{dx^i}\Gamma(x^i)=\frac{1}{(x^i-z)^2}\left[-(a+\delta)p^i+q_0(x^i-x_0^i)+((a+\delta)\pi^i-q_0)\Delta x^i\right].$$

Hence as long as the condition for q_0 in (4.53) is satisfied, $\frac{d}{dx^i}\Gamma(x^i)<0$, and Proposition 4.2.5 is proved. □

Remark 1. Given that $\Psi(x^{\min})$ is an upper bound of $\left\{\Psi(x^i),\ \forall i\in[N]\right\}$ by Proposition 4.2.5, for some pressure field candidate q_t, if the most demanding condition $q_t \geq \Psi(x^{\min})$ is satisfied, then we are guaranteed that all best responses under such q_t are going to have the at most single switching policies. Thus, this suggests that one can adopt a sequential search algorithm for fixed point pressure fields such that the successive candidates always produce single switching best responses by verifying that the condition $q_t^{(k)} \geq \Psi^{(k)}(x^{\min})$ is met at each iteration $k=1,2,\cdots$. This algorithm is discussed in the next section.

4.3 Numerical Solution of Advection Equation

It is recognized that the advection equation in (3.7) for each Θ_k is a balance equation and conservative. Therefore, if one wishes to conserve probability, a discretization based on finite volume methods is recommended in [LeVeque, 1992]. By discretizing the $x-t$ plane with time-step $k=\Delta t$ and x-step $h=\Delta x$, we define a grid with points (x_j,t_n) where $x_j=jh, j=0,1,2,\ldots$ and $t_n=nk, n=0,1,2,\ldots$. Recall that $m^{\Theta_k}(x,t)$ is defined to be the temperature pdf conditional on the initial temperature falling in the interval Θ_k. The solution $m^{\Theta_k}(x,t)$ to (3.7) should satisfy the integral form of the conservation laws.

Denote $g(m^{\Theta_k}) := m^{\Theta_k} v^{\bar{\theta_k}}$, where $v^{\bar{\theta_k}} = -a(x - \bar{\theta_k}) + b u_*^{\bar{\theta_k}}$, and $u_*^{\bar{\theta_k}}$ is to be computed from the single switching control law. Then we have

$$\begin{aligned}\int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_{n+1}) dx = &\int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_n) dx \\ &- \left[\int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j+1/2}), t) dt - \int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j-1/2}), t) dt \right].\end{aligned} \tag{4.54}$$

We define M_j^n as an approximation to the cell average of $m^{\Theta_k}(x, t_n)$ for $x \in [x_{j-1/2}, x_{j+1/2})$ at t_n, and the numerical flux $F(M_j^n, M_{j+1}^n)$ which is the average "probability current" through $x_{j+1/2}$ over the time interval $[t_n, t_{n+1}]$

$$\begin{aligned} M_j^n &\simeq \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_n) dx, \\ F(M_j^n, M_{j+1}^n) &\simeq \frac{1}{k} \int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j+1/2}), t)) dt. \end{aligned}$$

Accordingly the integral form of the conservation laws in (4.54) can be expressed as

$$M_j^{n+1} = M_j^n - \frac{k}{h}[F(M_j^n, M_{j+1}^n) - F(M_{j-1}^n, M_j^n)]. \tag{4.55}$$

Using a finite difference method such as Lax-Friedrichs or Lax-Wendroff (see [LeVeque, 1992]), we can numerically solve for M_j^n which is approximation to the solution $m^{\Theta_k}(x, t)$ on the defined grid. For example, if we use the Lax-Friedrichs scheme, which has the following form

$$M_j^{n+1} = \frac{1}{2}\left(M_{j+1}^n + M_{j-1}^n\right) - \frac{k}{2h}\left(g(M_{j+1}^n) - g(M_{j-1}^n)\right).$$

We can therefore write the expression $F(M_j^n, M_{j+1}^n)$ as

$$\begin{aligned} F(M_j^n, M_{j+1}^n) = &\frac{h}{2k}(M_j^n - M_{j+1}^n) \\ &+ \frac{1}{2}\left(g(M_j^n) + g(M_{j+1}^n)\right). \end{aligned} \tag{4.56}$$

Using initial density $M_0^0 = m^{\Theta_k}(x_0, 0)$, we can get the propagation of an approximated solution M_j^n on the defined grid points (x_j, t_n) from (4.55) and (4.56).

4.4 Equilibrium Solution by Iteration Approach

As the equilibrium solution is a fixed point of the MFG system, we can compute it using an iterative approach. We propose the following numerical algorithm where the single switching control policy is adopted.

Given initial mass density $m(x_0, 0)$, we first make an arbitrary guess of a family of $m^{\Theta_k}(x, t)^0$ as the fixed point. Given y, z, and $\bar{x}_0$, we first compute q_y^∞, and initialize the algorithm by applying q_y^∞ to the MF equation system in (3.10). We follow the single switching control policy and solve the set of advection equations to compute $\{m^{\Theta_k}(x, t)\}$, based on which we compute a new q_t candidate. For this new candidate, we verify that the sufficient condition in (4.19) holds by choosing a proper $k_p(t)$, such that all $\frac{dp^i}{dt} \leq 0$. We repeat the procedure until $m^{\Theta_k}(x, t)$ converges to a fixed point, from which we get the equilibrium solution $(m^*(x, t), \{u_*\})$.

Algorithm 1 Compute fixed point of the MFG system

Compute
q_y^∞
q_t^-
$\Psi(x_-^{\min})$
$k_c = \Psi(x_-^{\min}) - q_t^-$
Initialize
$k \leftarrow 0$
$m^{(k)}(x,0) \leftarrow m^0(x_0)$
$q_t^{(k)} \leftarrow q_y^\infty$
$q_{in} \leftarrow 0$
while $\left\| q_t^{(k)} - q_{in} \right\|_\infty > \epsilon$ **do**
  $q_{in} \leftarrow q_t^{(k)}$
  Compute $m^{(k+1)}(x,t)$, $q_t^{(k+1)}$
  for $i = 0 : (T/\Delta t) - 1$ **do**
    $k_p(i\Delta t) \leftarrow 0$
    while $q_t^{(k+1)}(i\Delta t) < \Psi(x^{\min}(i\Delta t))$ **do**
      $k_p(i\Delta t) \leftarrow k_p(i\Delta t) + \Delta_k$
      $q_t^{(k+1)}(i\Delta t) \leftarrow q_t^{(k+1)}(i\Delta t) + k_p(i\Delta t)$
      Compute $\Psi(x^{\min}(i\Delta t))$
    end while
    return $q_t^{(k+1)}(i\Delta t)$
  end for
  return $q_t^{(k+1)}$
  $k \leftarrow k + 1$
end while
return $m^K(x,t)$

Remark 2. We can always find a sufficiently large $k_p(t)$ such that the inequality $q_t = \left|\lambda \int_0^t (\bar{x} - y)dt\right| + k_p(t) \geq \Psi(x^{\min})$ holds for all time t. Suppose that at some $k_p(t)$, the above inequality is not satisfied. We can then keep increasing k_p up to k_c, then according to (4.37), we must have $\left|\lambda \int_0^t (\bar{x} - y)dt\right| + k_c(t) = (q_t - k_p) + k_c \geq \Psi(x^{\min})$. Therefore it is guaranteed that we can always find a $k_p(t) \leq k_c(t)$, $\forall t$ from the inner loop of the algorithm such that (4.19) holds. Recall that $k_c = \Psi(x_-^{\min}) - q_t^-$ which can be explicitly computed.

Remark 3. Given the monotone property of $\Psi(x^i)$ by Proposition 4.2.5, it is

sufficient to compare the q_t candidate and $\Psi(x^{\min})$ in the inner loop to check the validity of (4.19).

CHAPTER 5 NUMERICAL STUDY

In the numerical study, we make the following assumptions. The distribution of initial temperature is uniform between 18°C and 28°C, hence the initial mean temperature is 23°C. We wish to achieve a target mean temperature y of 22°C over a 4 hours horizon, and set parameter z in (3.4) to 17°C so every device will tend to decrease their temperature. Other parameters used in the simulation are as follows: $x_a = 0$°C, $a = -0.03$, $\delta = 0.001$, $b = 0.2$, $r = 1$, $q_0 = 200$, $\lambda = 40$. The constraint we impose is that the total control $(u^i + u^i_{free})$ must be always greater than zero, as it is inadmissible to have a negative control to cool down the space. Hence for u^i, the constraint is $u^i \geq -u^i_{free}$.

We first compute q_y^∞ by

$$\begin{aligned} q_y^\infty &= \left(\frac{a(a+\delta)r + b^2 q_0}{b^2}\right)\left(\frac{\bar{x}_0 - y}{y - z}\right), \\ &= \left(\frac{0.03(0.03 + 0.001) + 0.2^2 \cdot 200}{0.2^2}\right)\left(\frac{23 - 22}{22 - 17}\right) = 40.0046. \end{aligned}$$

When we apply q_y^∞ to the system, we can compute q_t^-, $\Psi(p_-^{\min}, x_-^{\min})$, and k_c, as plotted in Figure 5.1. Before we move on to compute the fixed point, we verify that in this experiment when q_y^∞ is applied, we do have $\frac{dp_-^i}{dt} \leq 0$ as illustrated in Figure 5.2. We can also verify that the condition in (4.53) for the monotone property of $\Psi(x_-^i)$ is satisfied, hence the upper and lower bounds of the family of $\Psi(x_-^i)$ are obtained when taking the lowest and highest initial temperatures respectively, and the results are demonstrated in 5.3.

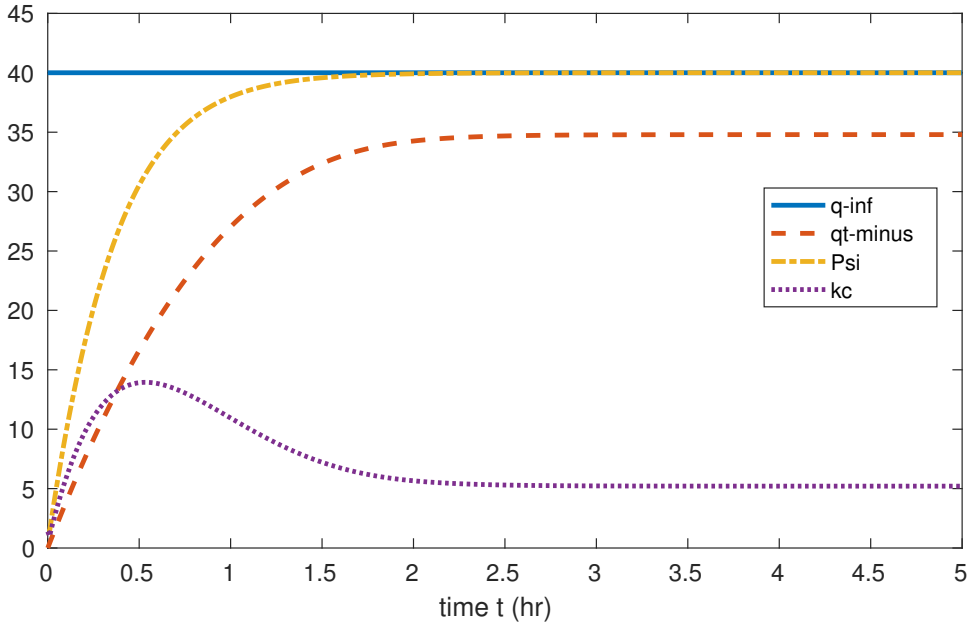

Figure 5.1 Plot of q_t^-, $\Psi(x_-^{\min})$, and k_c in the temperature decrease case

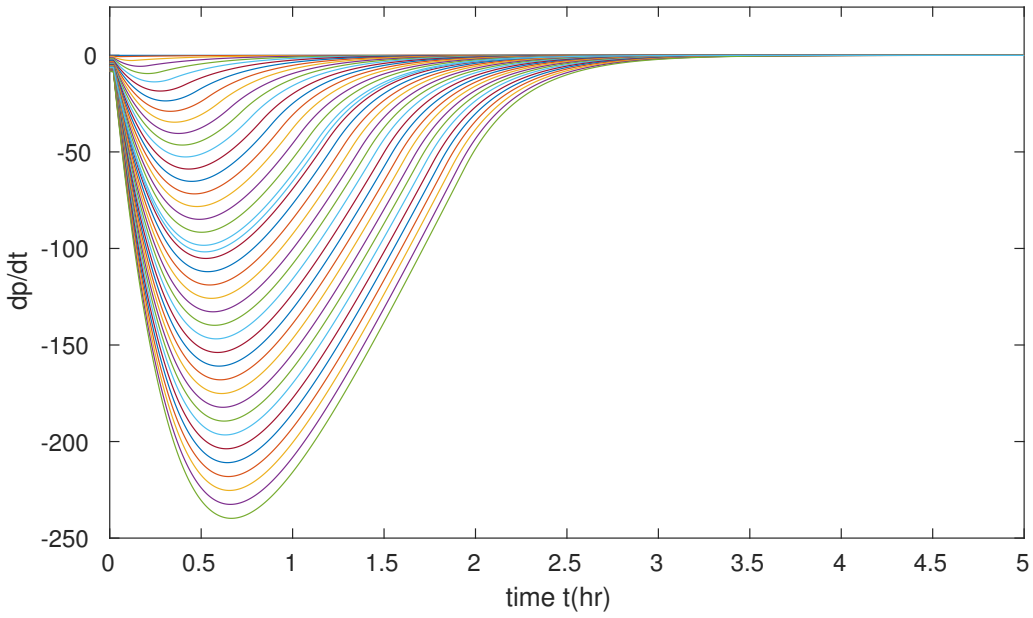

Figure 5.2 Plot of $\frac{dp_-^i}{dt} \leq 0$ in the temperature decrease case

Now suppose that we change some parameters such that the condition in

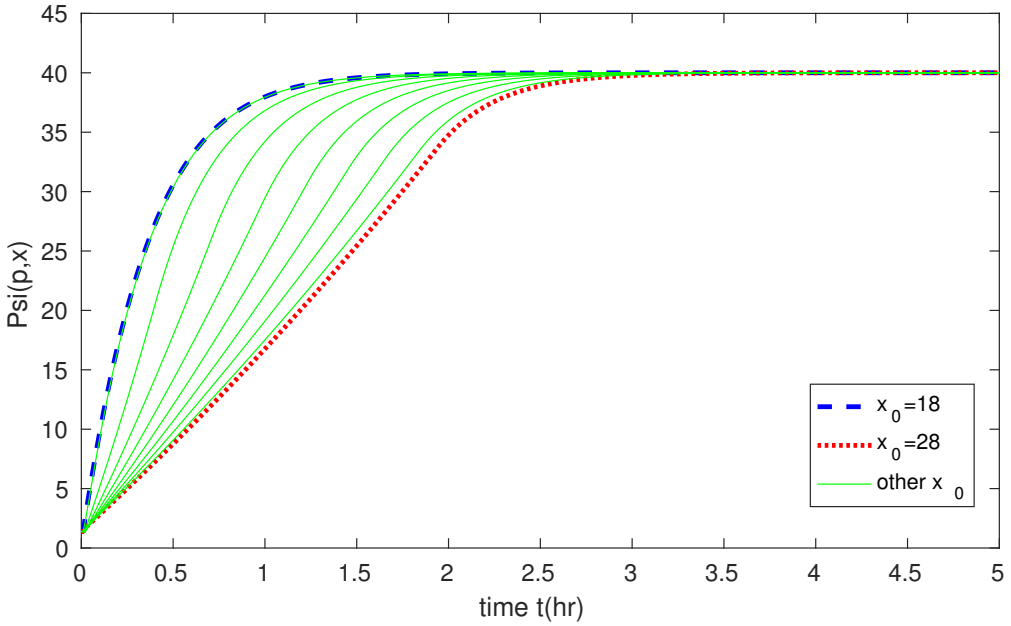

Figure 5.3 Plot of $\left\{\Psi(x_-^i)\right\}$ in the temperature decrease case

(4.53) is violated, for example, we choose $y' = 20°\text{C}$ and $z' = 17.5°\text{C}$. We have

$$\frac{\bar{x}_0 - y'}{y' - z'} = \frac{23 - 20}{20 - 17.5} = 1.2,$$

$$\left(\frac{a}{a+\delta}q_0\right) \Big/ \left(\frac{a(a+\delta)r}{b^2} + q_0\right) = 0.9676,$$

$$\frac{\bar{x}_0 - y'}{y' - z'} > \left(\frac{a}{a+\delta}q_0\right) \Big/ \left(\frac{a(a+\delta)r}{b^2} + q_0\right).$$

We plot the family of $\Psi(x_-^i)$ again under the new parameters, and we notice that the monotone property does not hold due to the violation of the condition in (4.53), as shown in Figure 5.4. Note that as the condition is only required when controls are saturated, the monotone property is therefore violated at the beginning of the control horizon, but resumes as controls are unsaturated.

We initialize the iteration process by using q_y^∞, and follow the numerical algorithm to find the fixed point solution. Figure 5.5 illustrates the evolution of q_t at each iteration which is constructed from the solutions to the MFG system until convergence. In all iterations, we verify that the condition in (4.19) is

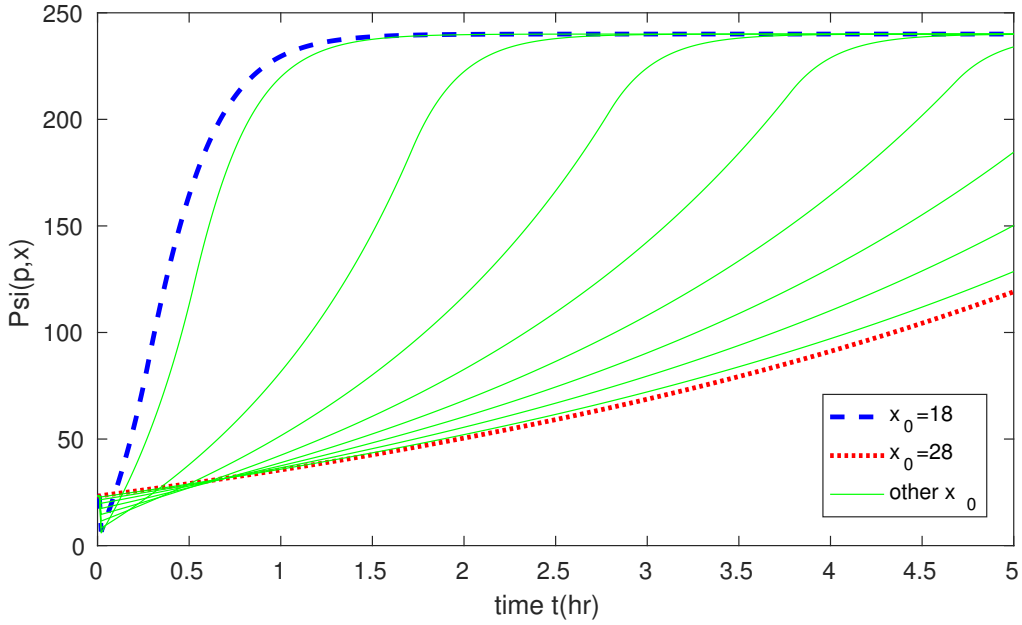

Figure 5.4 Plot of $\left\{\Psi(x^i_-)\right\}$ in the temperature decrease case when the monotone property is lost when $y' = 20°\text{C}$ and $z' = 17.5°\text{C}$

respected, by showing that $q_t - \Psi(x^{\min}) \geq 0$, $\forall t$ in Figure 5.6. Hence all costates $p^i(t)$ are monotonically decreasing such that the single switching control policy is adopted. Figure 5.7 illustrates the proper k_p term added to q_t in order to satisfy (4.19). Note that k_p obtained from the first iteration is in fact the k_c shown in Figure 5.1, and the required k_p in all remaining iterations are less than k_c, which is consistent with Remark 1.

Note that at each iteration, the q_t candidate is computed from the propagation of mass density $m(x,t)$ which is approximated by the family of $m^{\Theta_k}(x,t)$. Figure 5.8 shows the propagation of mass density $m^*(x,t)$ at the converged fixed point from the numerical algorithm. We start from an initial mass with uniform distribution, and at steady-state the mass shifts towards lower temperature and we arrive at another uniform distribution. Note that we are using the Lax-Friedrichs scheme, which is a first order finite volume method and the edges of the uniform distribution at steady-state are smeared off. By using a finer $x - t$ grid plane, we can reduce the smearing effects such that steady-state mass shapes more similar to a uniform distribution, but the CFL

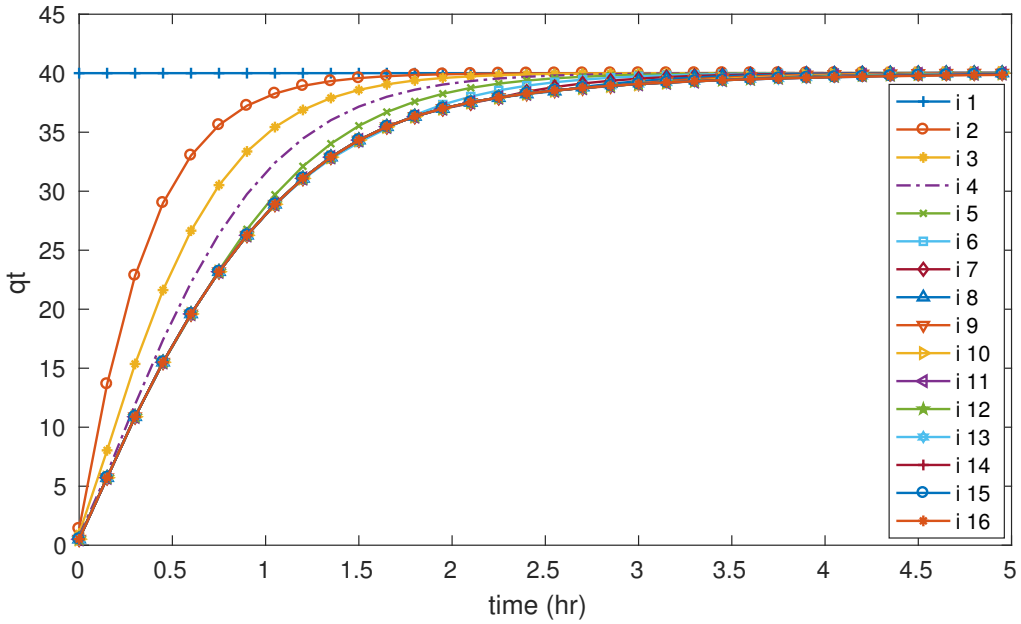

Figure 5.5 Evolution of q_t at each iteration from the numerical algorithm in the temperature decrease case

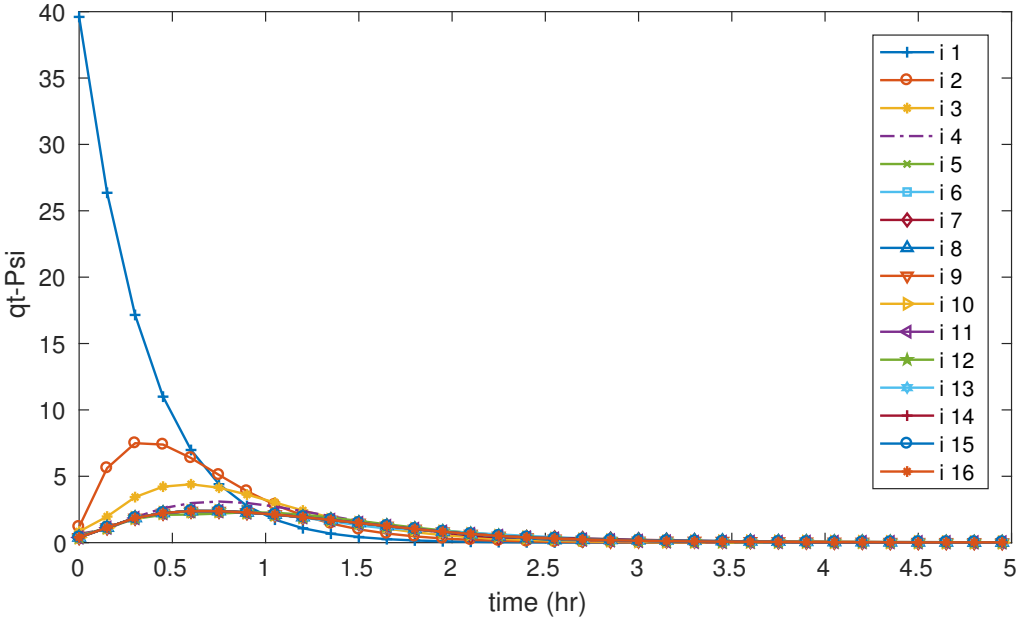

Figure 5.6 At each iteration $q_t - \Psi(x^{\min}) \geq 0$, $\forall t$ from the numerical algorithm

(Courant–Friedrichs–Lewy; see [LeVeque, 1992]) condition must always be satisfied in order to ensure stability of the numerical method. Specifically for the

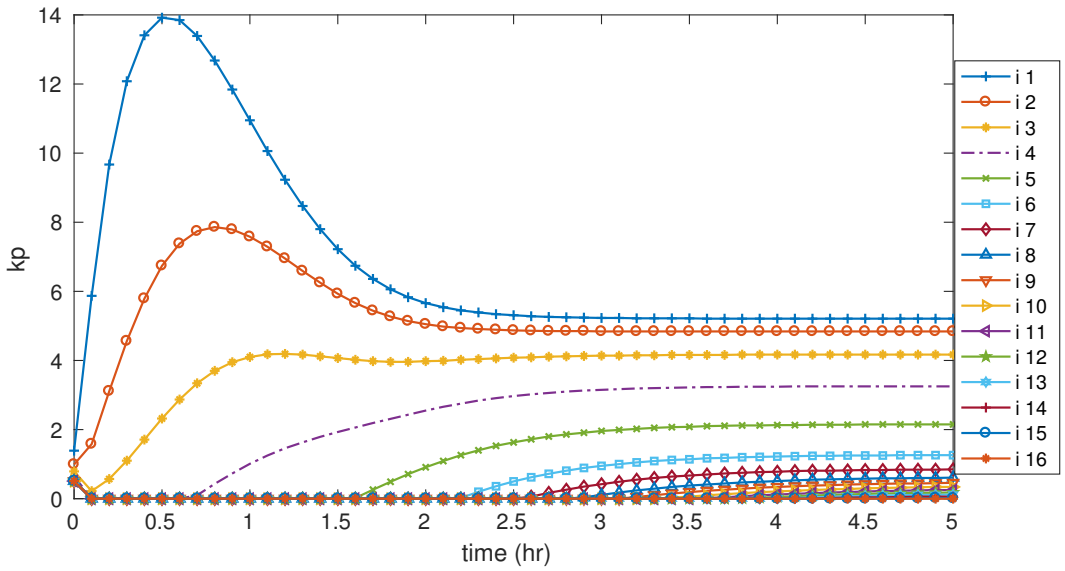

Figure 5.7 Evolution of k_p at each iteration to guarantee $\frac{dp^i}{dt} \leq 0$ in the temperature decrease case

first-order method, the time-step Δt and the x-step Δx must satisfy $\frac{\nu \Delta t}{\Delta x} \leq 1$, where ν is the magnitude of the drifting velocity of the mass. Also, note that the partition Θ_k of initial temperatures should be sufficiently refined in order to avoid significant ripples in the transient mass probability density.

In Figure 5.9 we take a few particle samples whose initial temperatures are randomly taken from the initial uniform distribution. We show the single switching behaviors of control trajectories $\{u_*^i\}$, and the corresponding state trajectories $\{x_*^i\}$ under the fixed point q_t. From the setup of the quadratic cost function in (3.4), the particles with higher initial temperatures feel more pressure to quickly decrease the temperature; however, due to the presence of the control constraint, their rates of temperature decrease are constrained at $t = 0$ (trajectories in red color). On the other hand, particles starting from lower initial temperatures feel less stress to decrease temperature, and their controls are not affected by the control constraint (trajectories in blue color). It is seen that the mean temperature of the population reaches the desired target temperature of 22°C at steady- state.

The derived model can also be used in the unconstrained case by simply

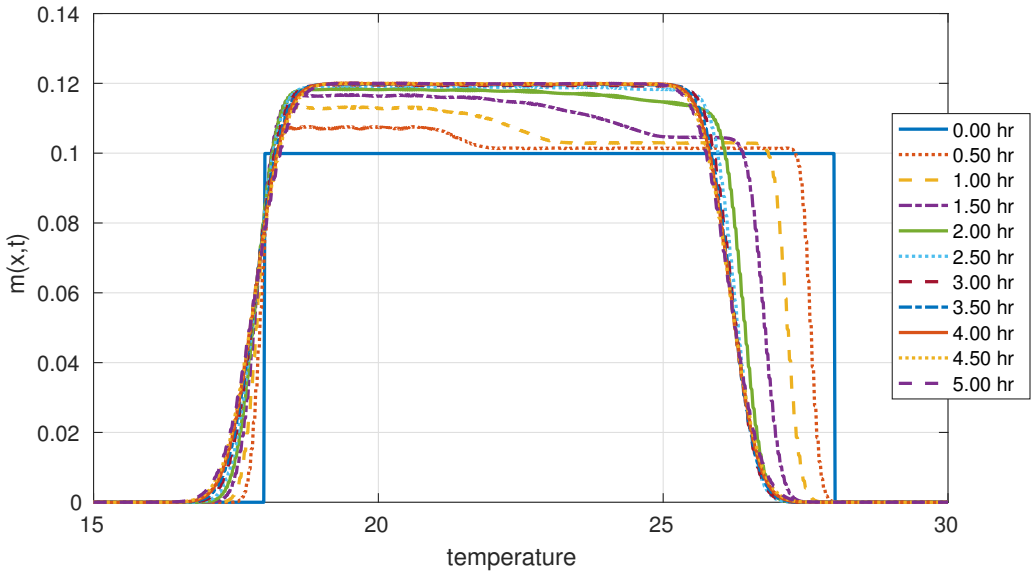

Figure 5.8 Evolution of $m(x,t)$ at the fixed point from the numerical algorithm in the temperature decrease case

removing the constraint and let $u^i \in \mathbb{R}, \ \forall i$ and set all switching time to $t = 0$. It is verified that under the same iteration approach we get an equilibrium solution identical to that in [Kizilkale et Malhamé, 2013].

In the next experiment, we wish to explore the differences in dynamics under constrained versus unconstrained cases. In particular we are interested in the mean temperature $\bar{x}(t)$ and trajectories of particle samples starting from the extreme initial temperatures. Figures 5.10 and 5.11 illustrate the comparison results. For particles starting from 28°C, as the imposed control constraints limit their rates of decrease since the beginning of the control horizon, the temperatures do not fall as fast as they do in the unconstrained case. As a consequence, q_t increases due to a larger error recorded between $\bar{x}$ and y, as compared to the unconstrained case at the same point in time. With a larger q_t, particles starting from 18°C are pressured to contribute more to reduce the global mean temperature. However, they do not feel enough stress to completely make up the gap caused by the constrained particles. Once controls become unsaturated, the constrained particles continue to decrease more than they would do in the unconstrained case, while the particles with low initial temperatures start to

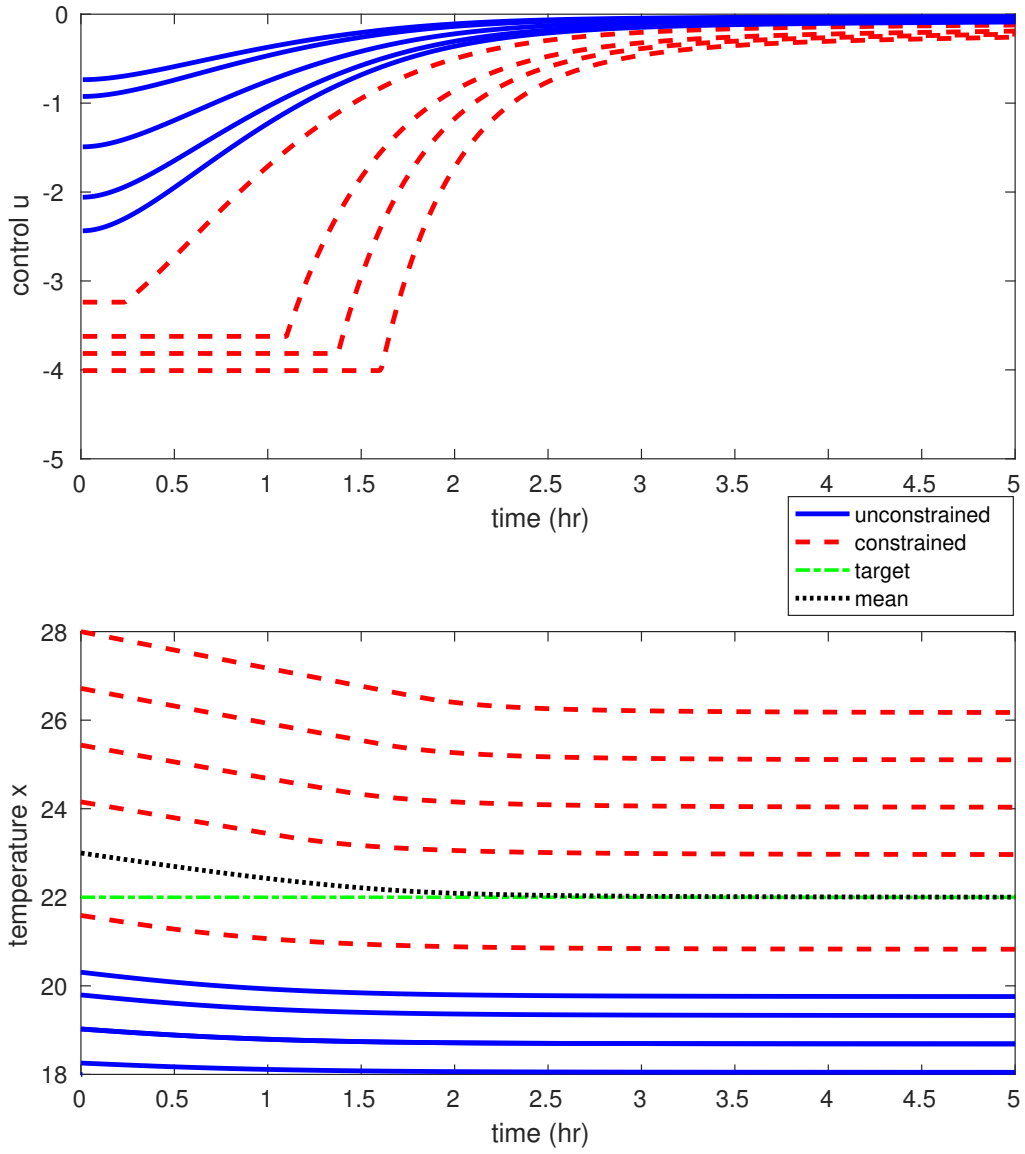

Figure 5.9 Trajectories of controls and states at the fixed point in the temperature decrease case

contribute less correspondingly. Eventually at steady-state, *all particles including the mean temperature have the same values as those in the unconstrained case*. This result is expected as all controls fall within the saturation limits at steady-state, hence controls are unconstrained.

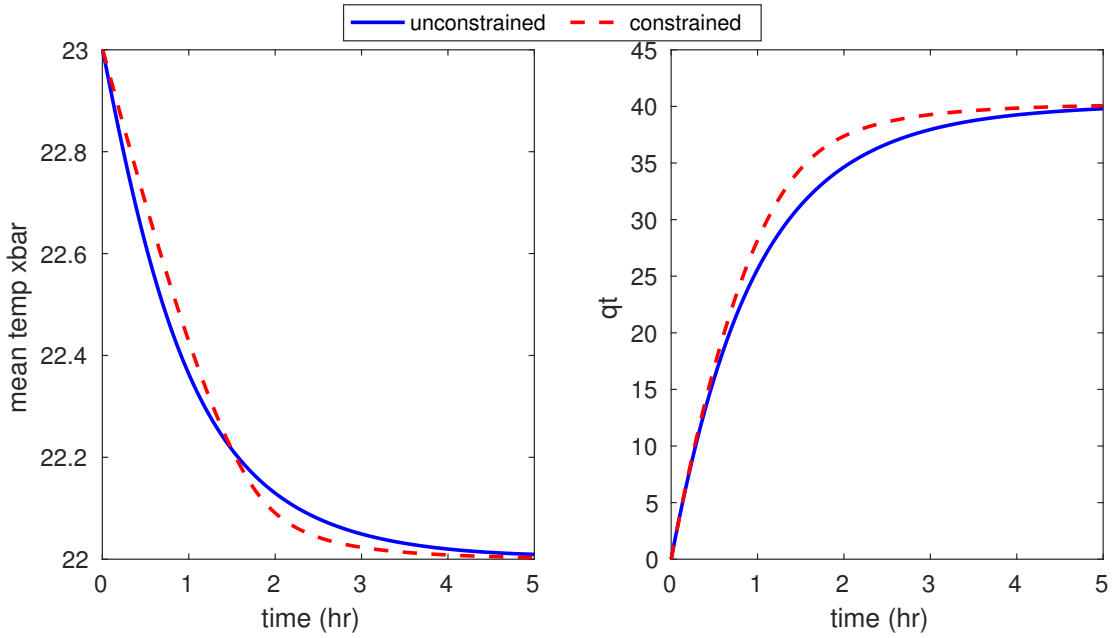

Figure 5.10 Comparison of mean temperature $\bar{x}(t)$ trajectories (**Left**) and comparison of fixed point q_t (**Right**) under constrained controls vs. unconstrained controls

Note that from the demonstrated results, if this control mechanism is to be put into practice, settling time may be a potential issue, i.e. the mean temperature is only reduced by 1°C within 5 hours. This is due to the fact that we rely on monotonicity of the costates in order to implement the single switching control policy, hence the pressure coefficient $q_t \leq q_y^\infty, \ \forall t$, and is overdamped in reaching the steady-state q_y^∞. In the next experiment, we wish to introduce some overshoot in the behavior of q_t by increasing the coefficient λ and allowing q_t to increase beyond q_y^∞ during transient. This will violate the monotonicity property of the costates, but as long as all costates $p^i(t)$ do not reach the critical level $p^i_{critical} = b^{-1} r u^i_{free}$ after control switches to unconstrained, we can still implement the single switching control policy. Figure 5.12 shows the results in this case.

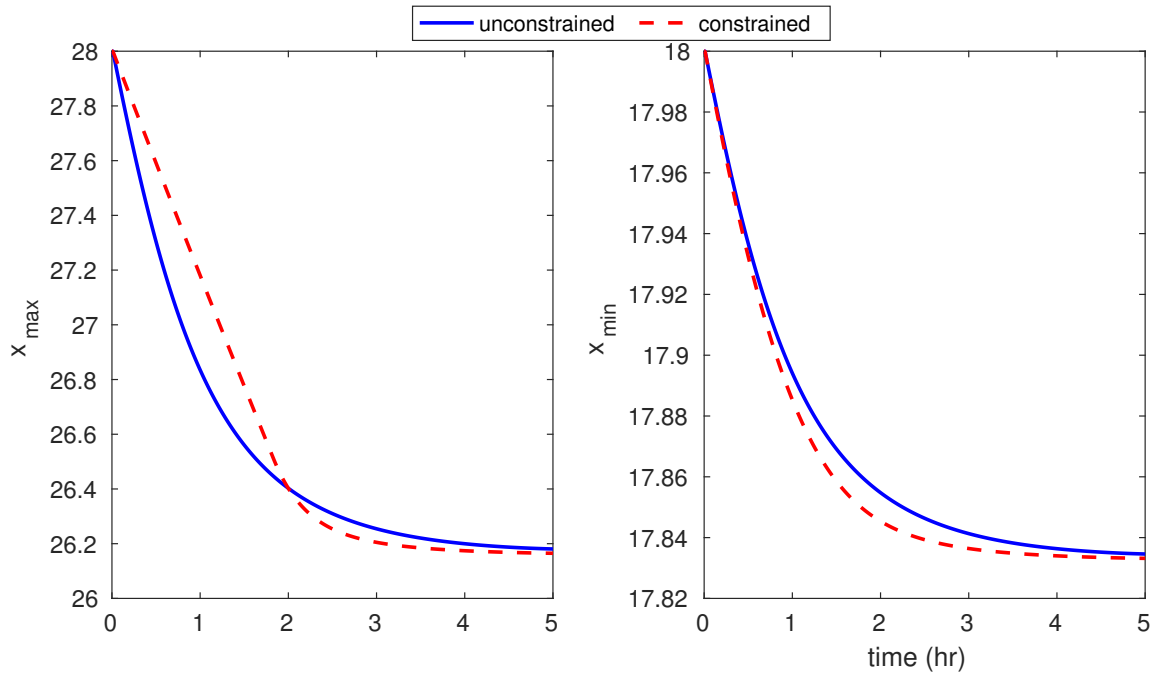

Figure 5.11 Comparison of temperature trajectories $x^{\max}(t)$ (**Left**) and temperature trajectories $x^{\min}(t)$ (**Right**) under constrained controls vs. unconstrained controls at fixed point

In the next set of experiments, we illustrate the case where all devices are required to heat when generation is at surplus. The distribution of the population's initial temperatures remain the same, and we ask the mean temperature to increase to $y = 24°\text{C}$. Accordingly, the parameter z is set to 28.5°C. Other parameters remain unchanged. The constraint we impose is that the total control $(u^i + u^i_{free})$ must be no greater than some $u_{\max} = 8$, hence for u^i, the constraint is $u^i \leq 8 - u^i_{free}$. In this case, we can still implement the single switching control policy by imposing that $\frac{dp^i}{dt} \geq 0$, which implies that at each iteration, we must have $q_t \geq \Psi(x^{\max})$. Note that in the heating case, the monotone property for $\left\{\Psi(x^i),\ \forall i\right\}$ holds as (4.53) is satisfied. Figure 5.13 shows the evolution of q_t to converge to a fixed point, and Figure 5.14 verifies that $q_t \geq \Psi(x^{\max})$. As in the temperature decrease case, we need a term k_p to keep the costate trajectories p^i monotonic in the heating case, and Figure 5.15 shows evolution of such k_p term until fixed point is reached.

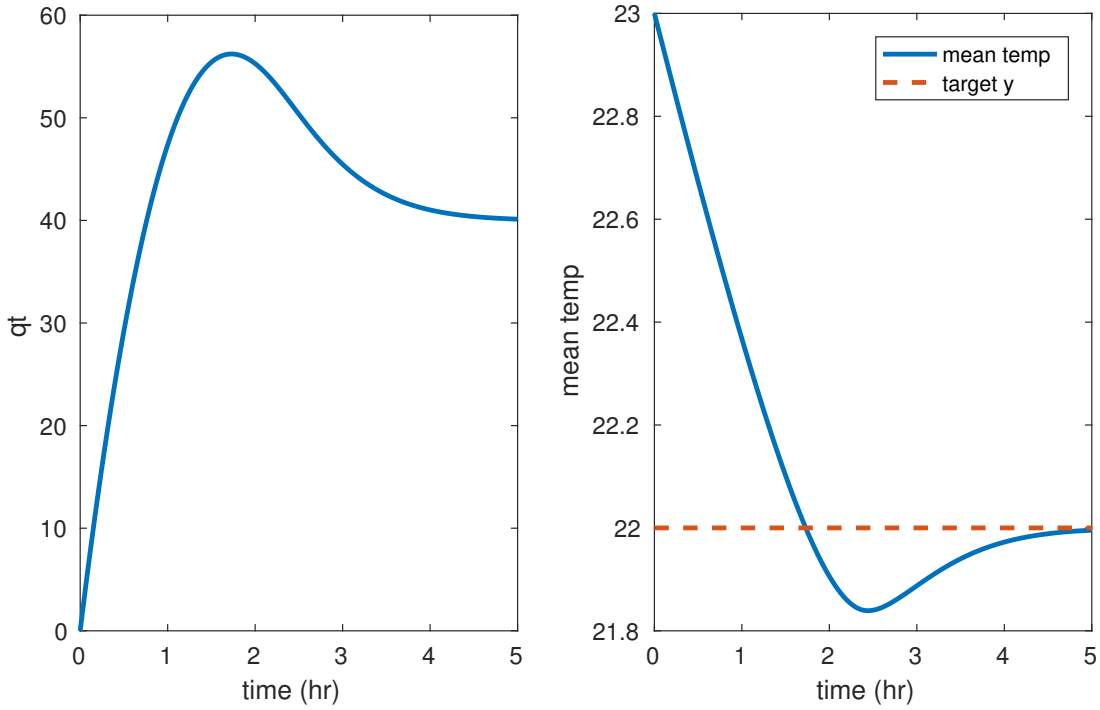

Figure 5.12 An underdamped q_t (**Left**) and the corresponding mean temperature $\bar{x}$ (**Right**) under single switching control policy

Figure 5.16 shows at the fixed point the control and state trajectories of particles whose initial temperatures are randomly chosen from the initial mass distribution. It is seen that as the mean temperature moves towards y, all controls have the single switching behaviors. Contrary to the temperature decrease case, the particles with lower initial temperatures have their controls saturated at the beginning of the control horizon, while the controls of the particles with higher initial temperatures remain unconstrained at all times.

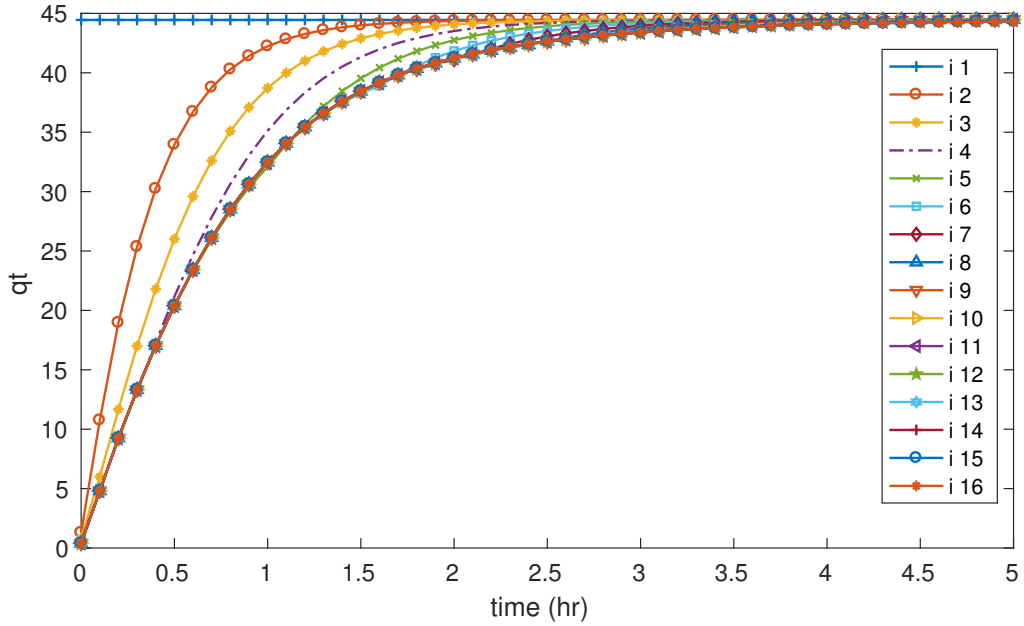

Figure 5.13 Evolution of q_t at each iteration from the numerical algorithm in the temperature increase case

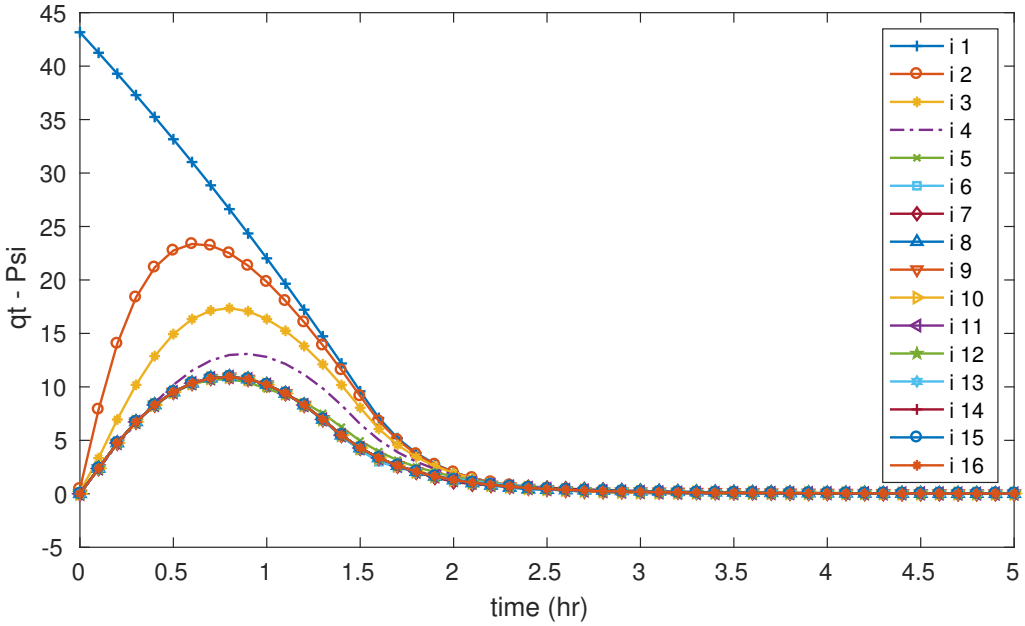

Figure 5.14 At each iteration $q_t - \Psi(x^{\max}) \geq 0$, $\forall t$ from the numerical algorithm in the temperature increase case

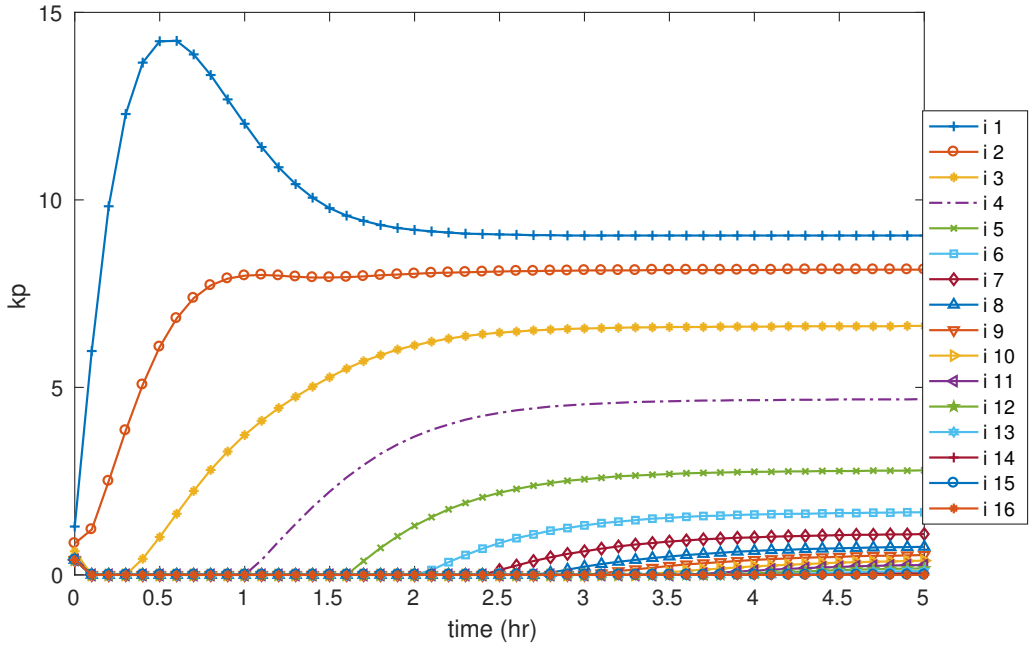

Figure 5.15 Evolution of k_p at each iteration to guarantee $\frac{dp^i}{dt} \geq 0$ in the temperature increase case

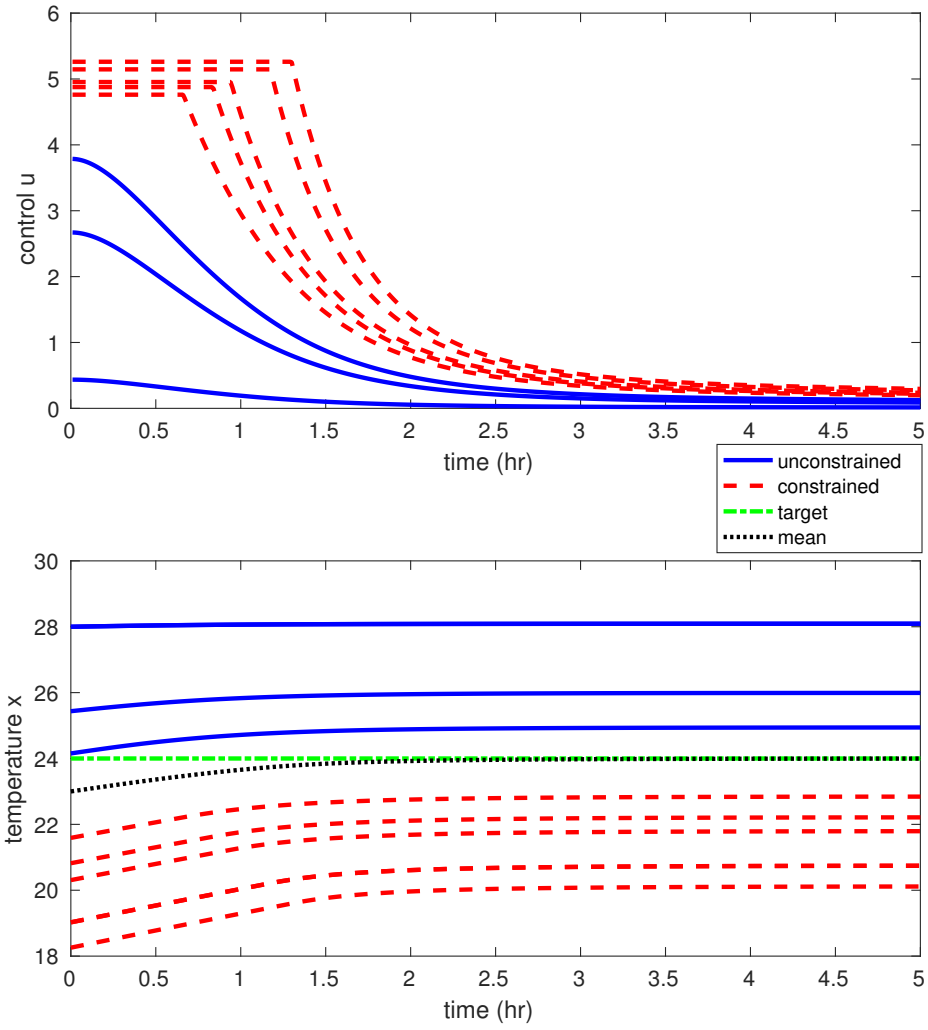

Figure 5.16 Trajectories of controls and states at fixed point in the temperature increase case

CHAPTER 6 CONCLUSION

In this thesis, we present a decentralized control mechanism to utilize a large number of electric space heaters as energy storage devices to mitigate the variability of renewable power generation, by letting the mean temperature of the device population follow a set temperature target, depending on the power generation status. The control for each device is generated locally and must respect certain control constraints. The control mechanism is based on the fixed point solution of the mean field formulation of the model, and we propose a numerical algorithm to compute an approximation of the fixed point where the desired mean temperature is reached. The non-linearity of the optimal control law under control constraints makes the computation of the MFG fixed point more challenging than in the linear quadratic case, requiring the solution of coupled PDE's, which moreover depends on the initial conditions of the devices. We describe an optimal control policy with single switching behaviors which is easy to implement, and we propose the sufficient condition for such control policy. The latter condition becomes the basis of an iterative numerical search algorithm for a fixed point characterizing a Nash equilibrium amongst controlled devices. We verify that the steady-states in the constrained case are identical to those in the unconstrained case.

In order to physically implement the developed control architecture, there are some practical considerations. For example, an infrastructure is required to facilitate communications between devices and the central authority, such that enough samples of initial conditions of devices can be read to form an estimate of their empirical temperature distribution, and the central authority can transmit the target y and the computed fixed point q_t to all devices. Also, given that all optimal controls to each device need to be generated locally, a controller must be physically installed on the device. While we may partially leverage the wide installation and use of smart meters in today's utility network, certain privacy concerns may also be raised for such implementation. Also, as seen from the last experiment of the numerical examples, the single switching control policy

based on the monotonicity condition of costates is slow in settling at the desired mean temperature target. In practice, a faster system response is required in order to efficiently overcome the variability issue of the renewable generation.

For future work, the existence of a fixed point remains to be proved for the constrained MFG equation system under a single switching control policy. We can start from the fact that under the sufficient condition in (4.37), the set of q_t must be bounded within q_t^- and q_y^∞, hence we can show that the set of q_t is non-empty, closed, and convex. We can then show that a fixed point exists by arguing that the set of q_t satisfies the conditions of Schauder's Fixed Point Theorem. To address the issue of slow response, we may modify the formulation of q_t as well as the sufficient condition in order to allow a more underdamped and oscillatory trajectory, but the control saturation can only happen during the first oscillation such that a single switching control policy can still be implemented. Finally, we wish to extend the analysis to the stochastic case where a noise process is included in (3.3), and to some multidimensional model where insulation and temperatures inside walls are considered.

REFERENCES

Bagagiolo, Fabio and Bauso, Dario (2013). Mean-Field Games and Dynamic Demand Management in Power Grids. *Dynamic Games and Applications*, *4*(2), 155–176.

A. Bemporad and M. Morari and V. Dua and E. N. Pistikopoulos (2000). The explicit solution of model predictive control via multiparametric quadratic programming. *American Control Conference, 2000. Proceedings of the 2000*. vol. 2, 872–876 vol.2.

Bryson, Arthur Earl (1975). *Applied optimal control: optimization, estimation and control*. CRC Press.

Callaway, Duncan S (2009). Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. *Energy Conversion and Management*, *50*(5), 1389–1400.

Cardaliaguet, Pierre (2010). Notes on mean field games. Rapport technique.

Ericson, Torgeir (2009). Direct load control of residential water heaters. *Energy Policy*, *37*(9), 3502–3512.

R. Goebel and M. Subbotin (2005). Continuous time constrained linear quadratic regulator - convex duality approach. *American Control Conference, 2005. Proceedings of the 2005*. 1401–1406 vol. 2.

Graichen, Knut and Kugi, Andreas and Petit, Nicolas and Chaplais, Francois (2010). Handling constraints in optimal control with saturation functions and system extension. *Systems & Control Letters*, *59*(11), 671–679.

Graichen, Knut and Petit, Nicolas (2008). Constructive methods for initialization and handling mixed state-input constraints in optimal control. *Journal Of Guidance, Control, and Dynamics*, *31*(5), 1334–1343.

S. Grammatico and B. Gentile and F. Parise and J. Lygeros (2015). A mean field control approach for demand side management of large populations of thermostatically controlled loads. *Control Conference (ECC), 2015 European*. 3548–3553.

S. Grammatico and F. Parise and M. Colombino and J. Lygeros (2016). Decentralized convergence to nash equilibria in constrained deterministic mean field control. *IEEE Transactions on Automatic Control*, *61*(11), 3315–3329.

P. Grieder and F. Borrelli and F. Torrisi and M. Morari (2003). Computation of the constrained infinite time linear quadratic regulator. *American Control Conference, 2003. Proceedings of the 2003*. vol. 6, 4711–4716 vol.6.

Gustafson, MW and Baylor, JS and Epstein, Gary (1993). Direct water heater load control-estimating program effectiveness using an engineering model. *IEEE transactions on power systems*, *8*(1), 137–143.

Hassan, MF and Boukas, EK (2008). Constrained linear quadratic regulator: continuous-time case. *Nonlinear Dynamics and Systems Theory*, *8*(1), 35–42.

Heemels, WPMH and Van Eijndhoven, SJL and Stoorvogel, AA (1998). Linear quadratic regulator problem with positive controls. *International Journal of Control*, *70*(4), 551–578.

M. Huang and P. E. Caines and R. P. Malhamé (2007). Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized epsilon-Nash Equilibria. *IEEE Transactions on Automatic Control*, *52*(9), 1560–1571.

Huang, Minyi and Malhamé, Roland P. and Caines, Peter E. (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, *6*(3), 221–252.

Donald E. Kirk (1970). *Optimal control theory: an introduction*. Prentice-Hall networks series. Prentice-Hall.

A. C. Kizilkale and R. P. Malhamé (2013). Mean field based control of power system dispersed energy storage devices for peak load relief. *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*. 4971–4976.

A. C. Kizilkale and R. P. Malhamé (2014a). A class of collective target tracking problems in energy systems: cooperative versus non-cooperative mean field control solutions. *2014 IEEE 53rd Annual Conference on Decision and Control (CDC)*. 3493–3498.

Kizilkale, Arman C and Malhamé, Roland P (2014b). Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads. *IFAC World Congress*.

Koch, Stephan and Zima, Marek and Andersson, Göran (2009). Potentials and applications of coordinated groups of thermal household appliances for power system control purposes. *2009 IEEE PES/IAS Conference on Sustainable Alternative Energy (SAE)*. IEEE, 1–8.

Laurent, JC and Malhamé, RP (1994). A physically-based computer model of aggregate electric water heating loads. *IEEE Transactions on Power systems*, *9*(3), 1209–1217.

LeVeque, Randall J (1992). *Numerical methods for conservation laws*, vol. 132. Springer.

F. Li and R. P. Malhamé and J. Le Ny (2016). Mean field game based control of dispersed energy storage devices with constrained inputs. *2016 IEEE 55th Conference on Decision and Control (CDC)*. 4861–4866.

Malhamé, Roland and Chong, Chee-Yee (1985). Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. *IEEE Transactions on Automatic Control*, *30*(9), 854–860.

D. Paccagnan and M. Kamgarpour and J. Lygeros (2015). On the range of feasible power trajectories for a population of thermostatically controlled loads. *2015 IEEE 54th Annual Conference on Decision and Control (CDC)*.

P. O. M. Scokaert and J. B. Rawlings (1998). Constrained linear quadratic regulation. *IEEE Transactions on Automatic Control*, *43*(8), 1163–1169.

M. M. Seron and J. A. De Dona and G. C. Goodwin (2000). Global analytical model predictive control with input constraints. *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*. vol. 1, 154–159 vol.1.

Sontag, Eduardo D (2013). *Mathematical control theory: deterministic finite dimensional systems*, vol. 6. Springer Science & Business Media.

Tarroja, Brian and Mueller, Fabian and Eichman, Joshua D and Samuelsen, Scott (2012). Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing. *Energy*, *42*(1), 546–562.

ANNEXE A BACKGROUND OF LQG MEAN FIELD GAMES

Suppose that we have a system of N non-uniform agents, where each agent i follows the following stochastic dynamics

$$dx_i = (a_i x_i + b_i u_i)dt + \sigma_i dw_i,\ t \geq 0 \tag{A.1}$$

where $1 \leq i \leq N$, x_i is the state of agent i, u_i is the control input (or strategy) of agent i, and w_t^i is a scalar Wiener process. The initial condition $x_i(0)$ is independent of w_t^i.

We consider a linear quadratic tracking problem by defining the following cost function,

$$J_i(u_i, x^*) = \mathbb{E}\int_0^\infty e^{-\rho t}\left\{(x_i - x^*)^2 + r u_i^2\right\} dt, \tag{A.2}$$

where $\rho,\ r > 0$, and x^* is a deterministic target.

The optimal control law to such LQG tracking problem to minimize $J_i(u_i, x^*)$ is

$$u_i^* = -br^{-1}(\Pi_i x_i + s_i), \tag{A.3}$$

where Π_i and s_i satisfy the following equations.

$$\begin{aligned} b^2 r^{-1}\Pi_i^2 - (2a_i - \rho)\Pi_i - 1 = 0, \\ -\frac{ds_i}{dt} = (a_i - \rho)s_i - b^2 r^{-1}\Pi_i s_i - x^*. \end{aligned} \tag{A.4}$$

Note that the equation for Π_i in (A.4) is an algebraic Riccati equation as the control horizon is infinite. Therefore we can explicitly compute $\Pi_i \geq 0$, and compute s_i in backwards. Therefore we the optimal control u_i^* in (A.3) is unique.

By substituting (A.3) into (A.1), we get

$$dx_i = (a_i - b^2 r^{-1}\Pi_i)x_i dt - b^2 r^{-1} s_i dt + \sigma_i dw_i. \tag{A.5}$$

By taking expectation on both sides of (A.5) and denoting $\bar{x}_i = \mathbb{E}[x_i]$, we get

$$\frac{d\bar{x}_i}{dt} = (a_i - b^2 r^{-1}\Pi_i)\bar{x}_i - b^2 r^{-1} s_i. \tag{A.6}$$

To consider the tracking problem in mean field game setup, we express x^* in terms of the population mean $\bar{x}^N = 1/N \sum_{i=1}^{N} \bar{x}_i$,

$$x^* = \gamma(\bar{x}^N + \eta), \tag{A.7}$$

where γ, $\eta > 0$. Given a sufficiently large population of N, $\bar{x}^N$ becomes deterministic, hence x^* is deterministic. Note that under this setup, all agents are weakly-coupled by the x^* term in the individual cost function.

We then can write the following mean field equation system:

$$-\frac{ds_a}{dt} = (a - \rho)s_a - b^2 r^{-1}\Pi_a s_a - x^*. \tag{A.8}$$

$$\frac{d\bar{x}_a}{dt} = (a - b^2 r^{-1}\Pi_a)\bar{x}_a - b^2 r^{-1} s_a. \tag{A.9}$$

$$\bar{x} = \int_{\mathcal{A}} \bar{x}_a dF(a), \tag{A.10}$$

$$x^* = \gamma(\bar{x} + \eta), \tag{A.11}$$

where $\Pi_a = rb^{-2}\left[a - \frac{\rho}{2} + \sqrt{\left(a - \frac{\rho}{2}\right)^2 + b^2 r^{-1}}\right]$. In the above equation system, the sequence $\mathcal{A} = \{a_i,\ 1 \leq i \leq N\}$ with an empirical distribution function $F(a_i)$. To simplify the notation, the coefficient a_i is replaced by a.

Denote $\beta_1(a) = -\frac{\rho}{2} + \sqrt{\left(a - \frac{\rho}{2}\right)^2 + b^2 r^{-1}}$, and $\beta_2(a) = \rho + \beta_1(a)$. Under the assumptions that:

1. $\beta_1(a) > 0,\ \ \forall a \in \mathcal{A}$, and $\int_{\mathcal{A}} M/(\beta_1(a)\beta_2(a))dF(a) < 1$, where $M = b^2\gamma r^{-1}$

2. $\mathbb{E}[x_i(0)] = 0, \ \forall i$

the equation system in (A.8)-(A.11) has a unique bounded solution for an infinite population, i.e. $N \to \infty$. Under this setup, given a posited x^*, all agents generate optimal controls according to (A.3), and for the solution to the equation system to be sustainable, the x^* must be replicated. In this sense, the unique solution is a fixed point Nash equilibrium of the equation system. The family of the optimal control policies is decentralized as they are computed locally by each individual agent as best responses to x^*.

As x^* is collectively constructed from the mean field of dynamics of agent i and all other agents of the population, we write $J_i(u_i, u_{-i})$ to indicate its dependence on control u_i and the set of controls of all other agents.

$$J_i(u_i, u_{-i}) = \mathbb{E} \int_0^\infty e^{-\rho t} \left\{ (x_i - \gamma(\bar{x} + \eta))^2 + r u_i^2 \right\} dt. \tag{A.12}$$

For a finite population $1 \leq N \leq \infty$, the decentralized family of optimal controls $\{u_i^*\}$ leads to an ε-Nash equilibrium, such that

$$J_i(u_i^*, u_{-i}^*) - \varepsilon \leq \inf_{u_i} J_i(u_i, u_{-i}^*) \leq J_i(u_i^*, u_{-i}^*). \tag{A.13}$$

As the number of population N increases, the equilibrium solution under the decentralized control policies gets closer to the Nash equilibrium obtained in the infinite population case.

ANNEXE B MATLAB CODES

B-1 Main program

The main program to compute the fixed point solution using the iteration algorithm.

```
1
2  %load the constant parameters required to run the numerical analysis
3  run('parameters');
4
5  %generate samples of particles from a uniform distribution
6  uniform_range=linspace(18,28,100)';
7  N=length(uniform_range);%
8
9  i=1;
10
11 %create variables
12 u=zeros(N,K);
13 x=zeros(N,K);
14 x_sol=zeros(N,K);
15 x_sol_cons=zeros(N,K);
16
17 qt_lb=zeros(N,K);
18 kp=zeros(1,K);
19 kp_his=[];
20 qt_lb_his=[];
21 indicator=0;
22 indicator_his=[];
23
24 mean(uniform_range)
25 size(unique(uniform_range))
26 pause
27
28
29 %initialize the state trajectories
30 x(:,1)=uniform_range;
```

```
31 x_sol(:,1)=x(:,1);
32 x_sol_cons(:,1)=x(:,1);
33
34 %compute u_free
35 u_offset=a*(xa*ones(N,1)-x(:,1))/b;
36
37 %compute qy^inf
38 qt_inf=((-a)*(-a+delta)*r+q0*b^2)/b^2*(mean(x(:,1))-y(1))/(y(1)-z);
39
40 %load the qt_minus and kc computed by using qy_inf
41 load 'qt_minus_kc_variable.mat';
42
43 tol=1e-2;
44 qt_new=qt_inf*ones(1,K);
45 qt_in=qt_new*0.8; %initial guess for qt
46 %kp(1)=kc(1);
47
48 %compute the constrained control for each individual sample
49 uc_min=u_min_total-u_offset;
50 uc_max=u_max_total-u_offset;
51
52 qt_his_i=qt_new;
53 qt_his=[];
54 xbar_his=[];
55
56 t_star=ones(1,N); % this is the index of t*
57 t_star_his=t(t_star);
58
59 %-----
60 %to construct the exact solution under constrained control
61 %-----
62 for k=2:K
63     %for temp decrease case
64     x_sol_cons(:,k)=x_sol_cons(:,1)+b/a*uc_min.*(exp(a*t(k))-1);
65 end
66 x(:,K)=x_sol_cons(:,K);
67 x(:,1)=x_sol_cons(:,1);
68
```

```
%The iterative approach to compute fixed point in terms of qt
while max(abs(qt_in-qt_new))>tol
    qt_his=[qt_his;qt_his_i];
    qt_in=qt_new;

    pi_theta_old=[];     %pi when control is not constrained
    s_theta_old=[];      %s when control is not constrained
    pi_theta=[];         %pi when control is constrained
    s_theta=[];          %s when control is constrained
    qt_lb=zeros(N,K);    %this is Psi under qt candidate in each
        iteration


    for n=1:N

        %solving Riccati from backwards for pi and s - to be used to
            find
        %switching time t*

        %generate pi(t) and s(t) for each agent from backwards under
        %non-constrained controls
        [t,pi_i,s_i]=pi_s_solution(qt_in,x(n,1));
        pi_theta_old=[pi_theta_old;pi_i];
        s_theta_old=[s_theta_old;s_i];


        %loop to find switching time t*
        kk=1;
        while kk<K
            if -b/r*(pi_theta_old(n,kk)*x_sol_cons(n,kk)+s_theta_old(n
                ,kk))>uc_min(n)
                break;
            end
            kk=kk+1;

        end

        t_star(n)=kk; % this is t*
```

```
104
105         %for t \in [0,t*) control constrained so state trajectory and
                control are known
106         x(n,1:t_star(n))=x_sol_cons(n,1:t_star(n));
107         u(n,1:t_star(n))=uc_min(n);
108
109         %solving Riccati from backwards for pi and s, using conditions
                at t*
110         %as terminal conditions
111
112         %generate piecewise pi and s for t \in [0,t*) and t \in [t*,T]
113         [t,pi_i,s_i]=pi_s_solution2(qt_in,x(n,1),t_star(n),uc_min(n));
114         pi_theta=[pi_theta;pi_i];
115         s_theta=[s_theta;s_i];
116
117         %to find state trajectory and optimal policy under
                unconstrained constrol
118         for k=t_star(n):K-1
119             u(n,k)=-b/r*(pi_theta_old(n,k)*x(n,k)+s_theta_old(n,k));
120             x(n,k+1)=x(n,k)+delta_t*(a*(x(n,k)-x(n,1))+b*u(n,k));
121         end
122
123       qt_lb(n,:)=((-a+delta)*(pi_theta(n,:).*x(n,:)+s_theta(n,:))+q0*(
              x(n,1)-x(n,:))./(x(n,:)-z);
124     end
125
126     t_star_his=[t_star_his;t(t_star)];
127     qt_lb_his=[qt_lb_his;qt_lb(1,:)];
128
129     %compute the costate trajectory
130     p=pi_theta.*x+s_theta;
131
132     x_bar=mean(x);
133     [x_bar,grid_range,mm]=propagation_of_mass_single_agent10v2(x,u,
            uniform_range);
134
135     %construct qt candidate before kp is computed
136     qt_new(1)=0;
```

```
137     for k=2:K
138         qt_new(k)=max(qt_minus(k),lambda*abs(trapz(0:delta_t:(k-1)*
                delta_t,(x_bar(1:k)-y(1:k)))));
139     end
140
141     xbar_his=[xbar_his;x_bar];
142
143
144     %verification of the sufficient condition for monotononic costate
145     %increase the term kp until the sufficient condition is met
146
147
148     %for k=1
149     k=1;
150     kp=zeros(1,K);
151     qt_new(1)=kp(1);
152     dpdt(:,1)=(-a+delta)*p(:,1)-qt_new(1)*(x(:,1)-z);
153     while max(dpdt(:,1))>0
154         kp(1)=kp(1)+0.01;
155         qt_new(1)=kp(1);
156
157         %recalculate the best response x and u under the new qt, as
                well as
158         %dp / dt
159         for n=1:N
160         [x_i_at_k,u_i_at_k,qt_lb_i_at_k,p_i_at_k]=comp_u_x_k(k,t_star(n
                ),x_sol_cons(n,:),uc_min(n),x(n,1),pi_theta_old(n,:),
                s_theta_old(n,:),pi_theta(n,:),s_theta(n,:));
161         p(n,k)=p_i_at_k;
162         x(n,k)=x_i_at_k;
163         u(n,k)=u_i_at_k;
164         end
165         dpdt(:,1)=(-a+delta)*p(:,1)-qt_new(1)*(x(:,1)-z);
166     end
167
168     %for k>1
169     for k=2:K
170         qt_new(1)=kp(1);
```

```
171         kp(k)=0;
172         dpdt(:,k)=(-a+delta)*p(:,k)-q0*(x(:,k)-x(:,1))-qt_new(k)*(x(:,k
                )-z);
173
174         while max(dpdt(:,k))>0
175         if qt_new(k)>=qt_inf
176              qt_new(k)=qt_inf;
177              break;
178         end
179         kp(k)=kp(k)+0.01;
180
181         for n=1:N
182         [x_i_at_k,u_i_at_k,qt_lb_i_at_k,p_i_at_k]=comp_u_x_k(k,t_star(n
                ),x_sol_cons(n,:),uc_min(n),x(n,1:k-1),pi_theta_old(n,:),
                s_theta_old(n,:),pi_theta(n,:),s_theta(n,:));
183         p(n,k)=p_i_at_k;
184         x(n,k)=x_i_at_k;
185         u(n,k)=u_i_at_k;
186         end
187         qt_new(k)=qt_new(k)+kp(k);
188
189         dpdt(:,k)=(-a+delta)*p(:,k)-q0*(x(:,k)-x(:,1))-qt_new(k)*(x(:,k
                )-z);
190         end
191
192         qt_his_i(k)=qt_new(k);
193         qt_his_i(1)=qt_new(1);
194         k
195
196     end
197         kp_his=[kp_his;kp];
198
199     %check that dpdt<0
200     dpdt=(-a+delta)*p-q0*(x-x(:,1)*ones(1,K))-(ones(N,1)*qt_new).*(x-z)
            ;
201     if max(max(dpdt))>0
202          indicator=1;
203     else
```

```
204             indicator=0;
205         end
206         indicator_his=[indicator_his;indicator];
207 
208         max(abs(qt_in-qt_new))
209         u_total=u+u_offset*ones(1,K);
210 
211         plot(t,qt_new),drawnow
212         hold on
213     end
```

B-2 $\pi^i(t)$ and $s^i(t)$ in the unconstrained case

The program to compute the $\pi^i(t)$ and $s^i(t)$ trajectories in the unconstrained case.

```
function [t,pi_theta,s_theta]=pi_s_solution(qt,x_init)

run('parameters');
t=[0:delta_t:T];
K=length(qt);
x0=x_init;
pi_theta=zeros(1,K);
s_theta=zeros(1,K);
pi_cons=zeros(1,K);
s_cons=zeros(1,K);

%compute the boundary conditions assuming at steady-state
pi_inf=care((a-0.5*delta),b,qt(K)+q0,r);
s_inf=(a*x0*pi_inf+qt(K)*z+q0*x0)/(a-delta-b^2/r*pi_inf);

pi_theta(K)=pi_inf;
s_theta(K)=s_inf;


for k = K:-1:2
    pi_theta(k-1)=pi_theta(k)-delta_t*(-(2*a-delta)*pi_theta(k)+b^2*(1/
        r)*(pi_theta(k))^2-qt(k)-q0);
end
for k=K:-1:2
    s_theta(k-1)=s_theta(k)-delta_t*(-(a-delta-b^2*pi_theta(k)*(1/r))*
        s_theta(k)+a*x0*pi_theta(k)+qt(k)*z+q0*x0);
end
end
```

B-3 $\pi^i(t)$ and $s^i(t)$ in the constrained case

The program to compute the $\pi^i(t)$ and $s^i(t)$ trajectories in the constrained case given the control switch time t^{i*}.

```
function [t,pi_theta,s_theta]=pi_s_solution2(qt,x_init,min_i,uc_min)

run('parameters');
t=[0:delta_t:T];
K=length(qt);
x0=x_init;
pi_theta=zeros(1,K);
s_theta=zeros(1,K);

pi_inf=care((a-0.5*delta),b,qt(K)+q0,r);
s_inf=(a*x0*pi_inf+qt(K)*z+q0*x0)/(a-delta-b^2/r*pi_inf);

pi_theta(K)=pi_inf;
s_theta(K)=s_inf;

%pi and s follow the same trajectories as in the unconstrained case
    until
%the control switch time t*
k=1;
for k = K:-1:min_i+1
    pi_theta(k-1)=pi_theta(k)-delta_t*(-(2*a-delta)*pi_theta(k)+b^2*(1/
        r)*(pi_theta(k))^2-qt(k)-q0);
end

for k=min_i:-1:2
    pi_theta(k-1)=pi_theta(k)-delta_t*(-(2*a-delta)*pi_theta(k)-qt(k)-
        q0);
end

%for t=[0,t*], pi and s to be computed from the equations in
    constrained
%case
k=1;
for k=K:-1:min_i+1
```

```
31      s_theta(k-1)=s_theta(k)-delta_t*(-(a-delta-b^2*pi_theta(k)*(1/r))*
            s_theta(k)+a*x0*pi_theta(k)+qt(k)*z+q0*x0);
32  end
33
34  for k=min_i:-1:2
35      s_theta(k-1)=s_theta(k)-delta_t*(-(a-delta)*s_theta(k)+a*x0*
            pi_theta(k)+qt(k)*z+q0*x0-b*pi_theta(k)*(uc_min));
36  end
37
38  end
```

B-4 Propagation of the mass density

The program to compute the propagation of the mass density and construct the mean temperature trajectory using a first order explicit finite-volume method.

```
function [xmean,grid_range,mm,m_profile]=
    propagation_of_mass_single_agent10v2(x,u,uniform_range)
run('parameters');

theta=20; %partition initial conditions into 40 intervals --> Theta
N=length(x(:,1));

subrange=[];
for i=1:theta
     subrange=[subrange uniform_range(1+(i-1)*N/theta:i*N/theta)];
end

x_theta=mean(subrange);
grid_size=1000;
grid_lower=10;
grid_upper=30;
pd=makedist('Uniform',18,28);
weight=pdf(pd,uniform_range);

grid_range=linspace(grid_lower,grid_upper,grid_size);%this is the
    global grid_range for everyone
m_profile=zeros(K,grid_size,theta);
delta_x=grid_range(2)-grid_range(1);

for i=1:theta;
x_subrange=x(1+(i-1)*N/theta:i*N/theta,:);
u_subrange=u(1+(i-1)*N/theta:i*N/theta,:);
f_subrange=a*(x_subrange-x_theta(i))+b*u_subrange;


dxdt=f_subrange; %this is size N/theta x K
dxdt(:,1)=dxdt(:,2);

m_j_n=zeros(K,grid_size,N/theta); %for single subrange with N/theta
```

```
    agents
if i==theta
    pd=makedist('Uniform',min(subrange(:,i)),28);
else
    pd=makedist('Uniform',min(subrange(:,i)),min(subrange(:,i+1))); %
        initialize to small uniform densities with x_theta as midpoint
end
m_j_n(1,:,:)=repmat(min(weight(1),pdf(pd,grid_range)),[1,1,N/theta]); %
    initialize all m_j_n to the same uniform density slice

%compute the propagation of the mass by first order explicit method
for n=1:K-1

    f_star=dxdt(:,n);
    f_j=grid_range;
    j=1:grid_size-1;
        f_jp1=f_j(j+1);
        f_jp1=[f_jp1 f_j(grid_size)];
        m_jp1=m_j_n(n,j+1,:);
        m_jp1=cat(2,m_jp1,m_j_n(n,grid_size,:));
    j=2:grid_size;
        f_jm1=f_j(j-1);
        f_jm1=[f_j(1) f_jm1];
        m_jm1=m_j_n(n,j-1,:);
        m_jm1=cat(2,m_j_n(n,1,:),m_jm1);

    j=2:grid_size-1;
        delta_m=m_j_n(n,j+1,:)-m_j_n(n,j-1,:);
        delta_m=[m_j_n(n,2,:) delta_m -m_j_n(n,grid_size-1,:)];
        delta_m2=m_j_n(n,j+1,:)+m_j_n(n,j-1,:)-2*m_j_n(n,j,:);
        delta_m2=[m_j_n(n,2,:)-2*m_j_n(n,1,:) delta_m2 m_j_n(n,
            grid_size-1,:)-2*m_j_n(n,grid_size,:)];

    for index=1:N/theta
        delta_f(:,:,index)=m_jp1(:,:,index).*f_jp1-m_jm1(:,:,index).*
            f_jm1;
        delta_f(:,:,index)=delta_m(:,:,index)*dxdt(index,n);
    end
```

```
66
67      m_j_n(n+1,:,:)=max(0,(m_j_n(n,:,:)-delta_t/(2*delta_x)*(delta_f))+
            delta_m2/2); %1st order
68
69  end
70
71  m_profile(:,:,i)=sum(m_j_n,3)/(N/theta);
72
73  end
74
75  mm=sum(m_profile,3);
76
77  xmean=zeros(size(t));
78  mm(1,:)=mm(1,:)/trapz(grid_range,mm(1,:));
79
80  %compute the mean of the mass from the propagation of mass density
81  xmean(1)=sum(grid_range.*mm(1,:))*(delta_x);
82  for k=2:K
83      mm(k,:)=mm(k,:)/trapz(grid_range,mm(k,:));
84      xmean(k)=max(xd,sum(grid_range.*mm(k,:))*(delta_x));
85  end
86  xmean=mean(x);
87  end
```